

THIS PAGE BLANK (USPTO)

Pa

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

BASE COUNT 670 a 1152 c 1076 g 722 t
ORIGIN

Alignment Scores:

Pred. No.: 2.46e-07 Length: 3620
Score: 90.00 Matches: 17
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 9

US-09-397-967-15 (1-17) x HSU31601 (1-3620)

OY 1 AAlaYsLeuLeuProLeuAspLysAspTyrTyValValArgGluProGly 17

Db 3006 GCTAGCTGCTGCGCTTGACAAAGACTACGTCGTCGCGAGCCAGGC 3056

RESULT 7

AX203081 AX203081 4064 bp DNA linear PAT 30-AUG-2001
DEFINITION Sequence 2 from Patent WO0152892.
ACCESSION AX203081
VERSION AX203081.1 GI:15392438

KEYWORDS

SOURCE

ORGANISM

human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 4064)
Jak/stat pathway inhibitors and the uses thereof

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

BASE COUNT

ORIGIN

Alignment Scores:

Pred. No.: 2.78e-07 Length: 4064
Score: 90.00 Matches: 17
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 6

US-09-397-967-15 (1-17) x AX203081 (1-4064)

OY 1 AAlaYsLeuLeuProLeuAspLysAspTyrTyValValArgGluProGly 17

Db 3006 GCTAGCTGCTGCGCTTGACAAAGACTACGTCGTCGCGAGCCAGGC 3056

RESULT 8

HSU09607 HSU09607 4064 bp mRNA linear PRI 09-JUL-1994
DEFINITION Human JAK family protein tyrosine kinase (JAK3) cDNA, complete cds.
ACCESSION U09607
VERSION U09607.1 GI:508730

KEYWORDS

SOURCE

ORGANISM

human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1124)
Nevanmure, M., McVicar, D.W., Johnston, J.A., Blake, T.B., Chen, Y.,
O'Shea, J., Doolittle, R., Schindler, C., Stiles, J.E., Ortaldo, J.R., and
Metcalf, C. Cloning of L-JAK, a Janus family protein-tyrosine kinase
expressed in natural killer cells and activated leukocytes
Proc. Natl. Acad. Sci. U.S.A. 91 (14), 6374-6378 (1994)
94397384

PUBMED 8022790
2 (bases 1 to 4064)

REFERENCE O'Shea, J. Direct Submission
TITLE Submitted (11-MAY-1994) John O'Shea, Leukocyte Cell Biology Section
JOURNAL LEI BRF, National Cancer Institute FCRDC, Bldg 560 Rm 3146 FCRDC,
Frederick, MD 21702, USA
Location/Qualifiers

FEATURES

source

gene

CDS

/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_type="natural killer cell, PHA activated T cells"
/gene="JAK3"
/gene="JAK3"
/gene="JAK3"
/codon_start=1
/product="JAK family protein tyrosine kinase"
/protein_id="AA19626.1"
/db_xref="GI:508731"

translation="MAPSEEMPLIPORSCSLSTPAGALHVLPPARPCPPORLSFS
EGDHLAEDLCVOAKASGILPYHSFALATEDISGPPSHFSVEDSTVLYLRI
REYTPMPELEKCHRGKRLDASLIDLPVLEHRAORSDLVSGLSLEBOG
ECTSLAVLDLARKAREOAPRPEGLKLTVAACLPRLSLDLOGLSFVTRAIIRTVR
RALPRVACQADHRSIMARKYIMDLERLDAGAEFTVGLPGALGHDGLRLKRVAGD
GGIATMGEQEVLPDPCDEPEYIDISIKOAPRYGAGEHRLVTRTDNQLAEPRG
LPLALFVLDVDFRLTDDSOHFEKQVAPPLLEVEAEOCGPITLDFAINKLKRG
GSRPGSYVLRSPDPSFLITVCNPNRGPVAVIRDLNLSIDVYELSDPFGALAP
RDLNMGACQLYACODPTFEERIKYISGLQNGRESVYLCDRPLAHTGALVANKO
LQNSGPDQDRDRETOIKALHSDITVTRGVSTGPGRLVNEYLPSCGLRDL
ORRRARLDASRLLYSSQICKMEYIGSRGVRRDLAANRILVESAHVKIDFGSLAK
LPLDLDYVVRPEPGSPITFWAPESLSNIFSRSDVSGFVGLYELFTYCDKSCSP
SAFLRMGCERDVPALCRLLLEEGORLPPAPACPAVHELMKLCMAPSPDRDRSF
SALGPOLDMLMSGRSCETHAFAPHEGHHLSFS"

BASE COUNT 746 a 1292 c 1168 g 858 t
ORIGIN

Alignment Scores:

Pred. No.: 2.78e-07 Length: 4064
Score: 90.00 Matches: 17
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 9

US-09-397-967-15 (1-17) x HSU09607 (1-4064)

OY 1 AAlaYsLeuLeuProLeuAspLysAspTyrTyValValArgGluProGly 17

Db 3006 GCTAGCTGCTGCGCTTGACAAAGACTACGTCGTCGCGAGCCAGGC 3056

RESULT 9

HSU70065 HSU70065 13562 bp DNA linear PRI 27-JAN-1997
DEFINITION Human JAK3 gene, complete cds.
ACCESSION U70065
VERSION U70065.1 GI:1800224

KEYWORDS

SOURCE

ORGANISM

human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 13562)
Riedy, M.C., Dutra, A.S., Blake, T.B., Modi, W., Lal, B.K., Davis, J.,
Bosse, A., O'Shea, J.J., and Johnston, J.A.
Genomic sequence, organization, and chromosomal localization of

JOURNAL
MODLINE
PUBLINE
REFERENCE
06

THIS PAGE BLANK (USPTO)

E08798 165 bp RNA linear PAT 29-SEP-1997
 LOCUS E08798
 DEFINITION cDNA encoding novel human tyrosine kinase.
 ACCESSION E08798
 VERSION E08798.1 GI:2176910
 KEYWORDS JP 1995059569-A/1.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1. (bases 1 to 165)
 Sakano, S.
 NOVEL TYROSINE KINASE PEPTIDE AND DNA CODING THE SAME
 TITLE Patent: JP 1995059569-A 1 07-MAR-1995.
 JOURNAL ASHIT CHEM IND CO LTD
 OS Homo sapiens (human)
 PN JP 1995059569-A/1
 PE 07-MAR-1995
 CF 25-AUG-1993 JP 1993210404
 SAKANO SEIJI

FEATURES
 source 1. 165
 FT source /organism="Homo sapiens"
 FT /cell_type="megakaryoblast"
 FT /cell_line="U-7"
 FT mat_peptide 1. 165
 FT /product="novel human tyrosine kinase"
 FT location/Qualifiers
 1. 165
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 36 a 56 c 39 g 34 t

BASE COUNT 36 a 56 c 39 g 34 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 9.04e-09 Length: 165
 Score: 90.00 Matches: 17
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 Gaps: 0

US-09-397-967-15 (1-17) x E08798 (1-165)
 OY 1 AlalysleuLeuProleuAspTyTyValValargGluProgly 17
 Db 52 GCTAAGCTGCTGCCGCTTGACAAAGACTACTAGCTGTCGCGAGCCAGGC 102

RESULT 2 HS08340 198 bp mRNA linear PAT 21-APR-1994
 LOCUS HS08340
 DEFINITION Human clone NTK16 tyrosine kinase mRNA, partial cds.
 ACCESSION U08340
 VERSION U08340.1 GI:473879
 KEYWORDS human.
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1. (bases 1 to 198)
 Fuortes, M.
 Tumor Necrosis Factor and Adhesion: Joint Control of Neutrophil
 TITLE Activation
 JOURNAL Thesis (1994) Cell Biology and Anatomy, Cornell University,
 Graduate School of Medical Sciences
 2 (bases 1 to 198)
 Fuortes, M.

TITLE Direct Submission
 JOURNAL Submitted (05-APR-1994) Michele Fuortes, Cell Biology and
 Anatomy/Medicine, Cornell University Medical College, 1300 York
 Avenue, New York, NY 10021, USA
 FEATURES
 source 1. 198
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="NTK16"
 /cell_type="neutrophil"
 /tissue_type="blood"
 <1..>198
 /codon_start=1
 /evidence="not experimental"
 /product="tyrosine kinase"
 /protein_id="AA17743.1"
 /db_xref="GI:473880"
 /db_xref="GI:473879"
 /translation="DLAARNLVESEAHVIADEGLAKLLPLDKYVYVREPGSDIF
 WYAPESLSDNFSRQSDVMSFG"

BASE COUNT 41 a 66 c 52 g 39 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1.1e-08 Length: 198
 Score: 90.00 Matches: 17
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 Gaps: 0

US-09-397-967-15 (1-17) x HS08340 (1-198)

OY 1 AlalysleuLeuProleuAspTyTyValValargGluProgly 17-
 Db 67 GCTAAGCTGCTGCCGCTTGACAAAGACTACTAGCTGTCGCGAGCCAGGC (117)

RESULT 3 AX203080 448 bp DNA linear PAT 30-AUG-2001
 LOCUS AX203080
 DEFINITION Sequence 1 from Patent WO0152892.
 ACCESSION AX203080
 VERSION AX203080.1 GI:15392437
 KEYWORDS human.
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1. (bases 1 to 448)
 Vasilos, G.
 Jak/stat pathway inhibitors and the uses thereof
 TITLE Patent: WO 0152892-A 1 26-JUL-2001;
 JOURNAL GENZYME CORPORATION (US)
 FEATURES
 source 1. 448
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 79 a 150 c 127 g 92 t

BASE COUNT 79 a 150 c 127 g 92 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 2.63e-08 Length: 448
 Score: 90.00 Matches: 17
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 Gaps: 0

US-09-397-967-15 (1-17) x AX203080 (1-448)

OY 1 AlalysleuLeuProleuAspTyTyValValargGluProgly 17
 Db 306 GCTAAGCTGCTGCCGCTTGACAAAGACTACTAGCTGTCGCGAGCCAGGC 356

RESULT 4
 LOCUS AX203087
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

THIS PAGE BLANK (USPTO)

GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 28, 2003, 17:00:25 ; Search time 6099 seconds
(without alignments)
5244.136 Million cell updates/sec

Title: US-09-397-967-16
Perfect score: 1099
Sequence: 1 MAPPEREPLIPORSCSLSS.....RPAFATLSPQLPLMGRG 1099

Scoring table:
OLIGO
Xgapop 60.0, Ygapext 60.0
Ygapop 60.0, Ygapext 60.0
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Word size: 1

Total number of hits satisfying chosen parameters: 4106490

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 150 summaries

Command line parameters:
-MODEL-frame+ p2n.model -DEV-xlp
-Q/cgn2.1/USPTO_spool/US09397967/runat_26042003.182926.11201/app.query.fasta.1.1287
-DB-GenEmbl -QFMT-fsstop -SUFFIX-oligo.rge -MINMATCH=0.1 -LOOPT=0 -LOOPEXT=0
-UNITS-bits -START=1 -END=1 -MATRIX-oligo -TRANS-human+0.cdi -LIST=150
-DOCALLIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=100 -MODE-LOCAL -OUTFMT=ptc
-NORM=ext -HAPSP52E=500 -MINLEN=0 -MAXLEN=2000000000
-USER-US09397967@cgn2.1.1.4675-efunat_26042003.182926.11201 -NCPU=6 -ICPU=3
-NO_XLPHY -NO_MMAPP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEVELTIMEOUT=120 -WARN_TIMEOUT=30 -THREDS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEX=7

Database: GenEmbl: *
1: gb_da: *
2: gb_hg: *
3: gb_in: *
4: gb_ov: *
5: gb_ov: *
6: gb_ov: *
7: gb_ov: *
8: gb_ov: *
9: gb_ov: *
10: gb_ov: *
11: gb_ov: *
12: gb_ov: *
13: gb_ov: *
14: gb_ov: *
15: gb_ov: *
16: gb_ov: *
17: gb_ov: *
18: gb_ov: *
19: gb_ov: *
20: gb_ov: *
21: gb_ov: *
22: gb_ov: *
23: gb_ov: *
24: gb_ov: *
25: gb_ov: *
26: gb_ov: *
27: gb_ov: *
28: gb_ov: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1037	94.4	3723	10	MSPTKIN
2	221	20.1	3528	10	MSJAK3A
3	180	16.4	4016	10	MSJAK3H
4	65	5.9	3778	10	RATJAK3
5	62	5.6	778	6	AX203087
6	62	5.6	2914	6	AX203084
7	62	5.6	3620	6	AX203084
8	62	5.6	4064	6	AX203081
9	62	5.6	4064	9	HSU09607
10	62	5.6	10326	10	MMU071201
11	62	5.6	13862	9	HSU70065
12	62	5.6	218208	2	AC073700
13	62	5.6	223734	2	AC073750
14	51	4.6	795	6	AX203086
15	50	4.5	159623	2	AC103459
16	40	3.6	448	6	AX203080
17	39	3.5	21323	9	AF513860
18	39	3.5	41006	9	AC007201
19	37	3.4	198	9	HSU008340
20	31	2.8	165	6	EB08798
21	30	2.7	3644	5	AF034576
22	29	2.6	811	6	AX203088
23	29	2.6	159623	2	AC103459
24	28	2.5	2881	9	BC028068
25	18	1.6	109	6	AX203082
26	18	1.6	726	6	AX203085
27	17	1.5	495	9	AF227557
28	16	1.5	466	9	HSU57096
29	16	1.5	1967	5	DRE5691
30	16	1.5	3435	6	AR129823
31	16	1.5	3495	6	AR048200
32	16	1.5	3495	6	AR068116
33	16	1.5	3495	6	AR071176
34	16	1.5	3495	6	AR177639
35	16	1.5	3495	6	162349
36	16	1.5	3495	6	187981
37	16	1.5	3500	9	AF001362
38	16	1.5	3629	6	AR136148
39	16	1.5	3629	6	AR143945
40	16	1.5	3629	6	192575
41	16	1.5	3731	10	MSPTK
42	16	1.5	3731	10	RNU13396
43	16	1.5	3739	5	AF148993
44	16	1.5	4042	4	AB036337
45	16	1.5	4161	9	AF005216
46	16	1.5	4387	5	DRE5690
47	16	1.5	4599	4	AB036336
48	16	1.5	4812	4	AB006011
49	16	1.5	5117	6	AX354568
50	16	1.5	5117	9	AF058925

51	16	1.5	66974	2	AC119228	AC119228 Mus muscu	124	1.1	2640	5	DRAJ5030	AJ005030 Dantio rer
52	16	1.5	85671	2	AC096324_3	Continuation (4 of	125	1.1	2720	9	HSEPHAT5	AF101169 Homo sapi
53	16	1.5	171146	9	AL161450	AL161450 Human DNA	126	1.1	2735	3	AB006565	AB006565 Ephydalia
54	15	1.4	4744	10	AF173032	AF173032 Mus muscu	127	1.1	2737	3	AB005305	AK095305 Homo sapi
55	15	1.4	7134	10	AF052607	AF052607 Mus muscu	128	1.1	2784	6	AX481390	AK095390 Sequence
56	15	1.4	14907	5	AF090382	AF090382 Tetradon	129	1.1	2820	5	GGECEK8A	Z19059 G. gallus Ce
57	15	1.4	172936	2	AC027154	AC027154 Mus muscu	130	1.1	2820	6	I15002	I15002 Sequence 5
58	14	1.3	8788	5	AF091238	AF091238 Tetradon	131	1.1	2839	9	AF037333	AF037333 Homo sapi
59	13	1.2	41	6	AR002620	AR002620 Sequence	132	1.1	2872	9	HSRETTM	X15786 Human ret-I
60	13	1.2	45	6	AR009670	AR009670 Sequence	133	1.1	2931	5	DRAJ5029	AJ005029 Dantio rer
61	13	1.2	45	6	AR002619	AR002619 Sequence	134	1.1	2934	10	AF131197	AF131197 Mus muscu
62	13	1.2	45	6	AR009669	AR009669 Sequence	135	1.1	2955	6	AX481387	AX481387 Sequence
63	13	1.2	1592	8	AY078944	AY078944 Arabidops	136	1.1	2955	6	AX481407	AX481407 Sequence
64	13	1.2	16977	8	TFU53213	TFU53213 Arabidops	137	1.1	2962	6	AR084933	AR084933 Sequence
65	13	1.2	107097	8	AC005315	AC005315 Arabidops	138	1.1	2962	6	AR084949	AR084949 Sequence
66	13	1.2	108393	8	AC005315	I44512 Sequence 11	139	1.1	2962	6	HUMPRTR	L36643 Homo sapien
67	12	1.1	151	6	I44512	AB012713 Anthocida	140	1.1	2962	6	AX481410	AX481410 Sequence
68	12	1.1	187	3	AB012713	M64614 Hydra vulg	141	1.1	2962	6	AX44405	AX44405 Sequence 2
69	12	1.1	201	3	HYDTRKIND	S75159 Homo sapien	142	1.1	2982	6	AR038071	AR038071 Sequence
70	12	1.1	206	6	H006874511	AX203083 Sequence	143	1.1	2982	6	AR172554	AR172554 Sequence
71	12	1.1	214	6	AX203083	AB038987 Gallus ga	144	1.1	2982	6	AX481391	AX481391 Sequence
72	12	1.1	302	5	AB038987	ML2257 Chicken DNA	145	1.1	2997	6	AX481409	AX481409 Sequence
73	12	1.1	373	5	CHRTYKDA	AR053289 Sequence	146	1.1	2997	6	AX481388	X91191 X laevis m
74	12	1.1	738	6	AR053289	AJ299010 Rattus no	147	1.1	3042	5	XLSEK1	AX481388 Sequence
75	12	1.1	790	10	RNC029010	U89380 Dantio rerio	148	1.1	3042	6	AX481388	AJ005028 Dantio rer
76	12	1.1	877	5	DRU89380	X76012 M. musculus	149	1.1	3055	5	DRAJ5026	L25890 Mus musculu
77	12	1.1	1128	10	MMSEK4	AB025542 Lampetra	150	1.1	3104	10	MUSNRK	
78	12	1.1	1153	5	AB025542	M64611 Hydra vulg						
79	12	1.1	1186	3	HYDTRKINA	D14717 Human RNA						
80	12	1.1	1225	3	HUMERK	X76010 M. musculus						
81	12	1.1	1294	10	MMSEK2	AR053290 Sequence						
82	12	1.1	1398	6	AR053290	AJ005028 Dantio rer						
83	12	1.1	1473	5	DRU5028	X13411 Rat mRNA fo						
84	12	1.1	1509	10	RNEK	AR053291 Sequence						
85	12	1.1	1521	6	AR053291	L43622 Xenopus lae						
86	12	1.1	1531	5	XLEPHRTK	AB025541 Epitretu						
87	12	1.1	1563	5	AB025541	X56348 H. sapiens u						
88	12	1.1	1568	9	HSURFRET	AB025538 Epitretu						
89	12	1.1	1582	5	AB025538	AB049595 Ephydalia						
90	12	1.1	1585	3	AB049595	U18084 Mus musculu						
91	12	1.1	1602	10	MMU18084	D37827 Homo sapien						
92	12	1.1	1618	9	HUMERKLP	AB049594 Ephydalia						
93	12	1.1	1619	3	AB049594	AL133666 Homo sapi						
94	12	1.1	1619	9	HSW801500	X76011 M. musculus						
95	12	1.1	1635	10	MMSEK3	U05210 Mus musculu						
96	12	1.1	1651	10	MMU05210	S77473 VNR-nonre						
97	12	1.1	1656	10	S77473	BC008655 Homo sapi						
98	12	1.1	1658	10	BC008655	L27738 Mus musculu						
99	12	1.1	1734	10	MUSNRK	S71669 Lsk-leukocy						
100	12	1.1	1744	9	S71669	AL137754 Homo sapi						
101	12	1.1	1804	9	HSW802388	X61601 S. lacustri						
102	12	1.1	1829	3	SLSRK1	L34542 Rattus norv						
103	12	1.1	1838	10	RATBATK	L03357 Homo sapien						
104	12	1.1	1867	9	HUMRETRI	DA5243 Mouse mRNA						
105	12	1.1	1911	10	MUSCTR	AR053292 Sequence						
106	12	1.1	1942	6	AR053292	E10345 cDNA encodi						
107	12	1.1	1942	6	E10345	X77278 H. sapiens H						
108	12	1.1	1968	9	HSHTYTR	AR084686 Sequence						
109	12	1.1	1987	6	AR084686	AR204704 Sequence						
110	12	1.1	1987	6	AR204704	L16974 Human tyros						
111	12	1.1	1987	9	HUMMATK	AR179662 Sequence						
112	12	1.1	2000	6	ARI79662	AK055395 Homo sapi						
113	12	1.1	2073	9	AK055395	X81466 M. musculus						
114	12	1.1	2112	10	MMEBK	AX015415 Sequence						
115	12	1.1	2170	6	AX015415	BC000114 Homo sapi						
116	12	1.1	2180	9	BC000114	BC003109 Homo sapi						
117	12	1.1	2180	9	BC003109	BC003072 Homo sapi						
118	12	1.1	2219	9	BC003072	AB006566 Ephydalia						
119	12	1.1	2318	3	AB006566	M27409 H. sapiens m						
120	12	1.1	2398	9	HSRTKEPH	M16029 Human ret m						
121	12	1.1	2485	9	HUMTYKRET	AJ236868 Xenopus 1						
122	12	1.1	2578	5	XLA236868	AJ236867 Xenopus 1						
123	12	1.1	2592	5	XLA236867							

ALIGNMENTS

RESULT 1

MUSPYTKIN

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

COMMENT

FEATURES

source

CDS

FEATURES

source

CDS

FEATURES

source

CDS

FEATURES

source

CDS

```
/translation="MAPSEETPLIPQSCSLSSSEAGALLVLPKPGPPORLSF  
EGDVLAEIDLCVRAKACALIPVYHSILAEIDEDSCPEPSHIECIDVPOVLYVL  
RFPYPMFGLFETCHRREGRLKRDLSAIIIDLIVLELHPORHSRSDLVSGSLPYGLSKBEG  
EFLSLAVDLAOMAREDOARQBELKTVSKACIPSLRDYIOGONVYTRIRRTYV  
LALLPCRLPGRPALMAKYILDERLHPATETFPVGLQAEFGLLRVADNRI  
PSSNDELQTFQFDFPEIADVSIINQARVPAGHRLVYIMQGHILAEFQLEPA  
LSFVALVDGYFRLICDSRHYECKEVPARLLEEADYCHPITLDEAIHKLKAGSLP  
GTYILARSPODYDFLLTACVOPFLGCGCLIROPSGSAFLVGLSDKSCSPAEEL  
AACNSCLRDGALALYTSCCAPRREKSNLIYVRRCNAPAGCSPSCALITLSF  
HTITPDSLEHENTHGSGFTKIFRGSREYVDGTHSEVLKMYDRNRHNCSELE  
AASMSQVSPHVLVHLGVCMAGDSINVDFVIGAITDMLTKRGHLSVSMKIQVTK  
QLAVALNLEDKGLPHGNVSAKRYLAREGGDGNPFIKILSDPVSPTVLSLEMLTOR  
IPWVAPCLOEAQTLCEADKMGFATTEVEFORGPAHITSLBPAKRLKYEEDGQLP  
ALKWTELGLAGITOCMAVDPGRRSPFRAILRLDNLMLITSDYELSDPTPGIPSPDELC  
VAGNOLACODDPAIEERHLKYSILGKNGSVELCYRDELDGNTGPIVAKLOHS  
VPDQDPOREIOIKALHSDFIYKRGVSGPROSLRLMEVLPESCLRDILQRRR  
GLTRDLRLFRMOJCKGMEYIGARCRHDAANILYVESAHYKINDFGIAKLIPLG  
KDYVVERPGSPILFWYAPESLSDNITSKQSDVSEFVLYELFTYCDKSCSPAEEL  
RMMGPEREGRPLRLLELLEGRRLPPPTCPTEVQELMDLQMAPEHDPAPFATLSP  
QLDPLMGRPG"  
13..3309  
/gene="jak3"  
/product="protein tyrosine kinase"  
/standard_name="Janus kinase, mjak3"  
/evidence=experimental  
3313..3723  
/gene="jak3"  
3723  
/gene="jak3"  
polyA_site  
BASE COUNT 711 a 1162 c 1079 g 771 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 0 Length: 3723  
Score: 1037.00 Matches: 1098  
Percent Similarity: 99.82% Conservative: 0  
Best Local Similarity: 99.82% Mismatches: 1  
Query Match: 94.36% Indels: 2  
DB: 10 Gaps: 0  
US-09-397-967-16 (1-1099) x MUSPTYKIN (1-3723)  
QY 1 MetAlaProProSerGluGluThrProLeuIleProGlnArgSerCysSerLeuSerSer 20  
DB 13 ATGGACCTCCCAAGTGGAGAGACACTTCATCCCTCAGCCCTCTGCGACCTTCATCC 72  
QY 21 SerGluAlaGlyAlaLeuHisValLeuLeuProProArgGlyProGlyProProGlnArg 40  
DB 73 TCAGAGGAGGAGGACCTTGATGCTCTCTCCCGGGGAGACTGGGCTCCCAAGCA 132  
QY 41 LeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysValArgAlaAlaIle 60  
DB 133 TTGTCAATCTCTTTGGGGAGCTACTTGCTGAGGATTTATGTGGAGCTGCCAAGGCC 192  
QY 61 CysGly-IleLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSerC 80  
DB 193 TGTGCG-TATCTGCGCTGTTATCATCTTCCTTTGCTGGCCACATGAGGACTTCTTG 251  
QY 80 SerProPheProProSerHisIlePheCysIleGluAspValAspThrGlnValLeuVal 100  
DB 252 CTGGTTTCCCAAGCCACATCTTCTGCATAGAGGAGCTGGAGACCTCAAGTCTGGTCA 311  
QY 100 rArgLeuArgPheTyrPheProAspTrpPheGlyLeuGluThrCysHisArgPheGlyLe 120  
DB 312 CAGGCTACGCTTTTATTTCCCTGACGTGTTGGGCTGGAGACATGTACACGGCTTGGGCT 371  
QY 120 uArgLysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGlnHisLeuPheAlaG 140  
DB 372 CGCAAGAGATTGTGACGAGGACATCTTGACTTATGACTTTTAAAGACATCTCTTGCTCA 431  
QY 140 HisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLysGlnGln 160  
DB 432 GCACCGCACTGAGCTGTGAGGCGCTCCGGGTGGGCTTAAAGCATGAGAGGAGG 491  
QY 160 YglIuPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGlnAlaGln 180  
DB 492 AGAGTCTCTGAGCTCGCGCGCTGCTGAGACTTGGCCAGATGGCTGTGAGAGCGGCCAGCG 551  
QY 180 gProGlyGluLeuLeuLysThrValSerTyrLysAlaCysLeuProProSerLeuArgAs 200  
DB 552 CCCAGAGAGGCTGCTGAAGAGCGCTACATTACAAAGCTGTCTGCCGCCACCTGCGCGCA 611  
QY 200 pValIleGlnIleGlnAsnPheValThrArgArgArgIleArgArgThrValValLeuAl 220  
DB 612 TGTGATCCAGGAGCCAGAACTTGTGTACAGCGAGCGCATCCGAGAGACGTGGTCTTGCG 671  
QY 220 aLeuLeuProCysGlyArgLeuProGlyArgProTyrAlaLeuMetAlaLysTyrIleLe 240  
DB 672 GCTGTGCGCGTGTGTGTGCTGCCCTGCCAGCGCGACCGTACAGCGCTCAATGGCCAGTATATCT 731  
QY 240 uAspLeuGluArgLeuHisIleProAlaIleThrThrGluThrPheArgValGlyLeuProG 260  
DB 732 GGACCTGGAGCGGCTACATCCAGCGCGCCACACCGAGACCTTCCGTGGGGCTTCCCGGG 791  
QY 260 YAlaGlnGluGluProGlyLeuLeuArgValAlaGlyAspAsnGlyIleProTrpSerSe 280  
DB 792 CCGCCAGAGAGAGCGCGGGCTTCTGCTGTGCGAGGGAGCAACGCAATCCCTTGAGAGCTC 851  
QY 280 rAsnAspLeuLeuPheGlnThrPheCysAspPheProGluIleValAspValSerIleAs 300  
DB 852 GAACGACGAGCTTTCAGACCTTGTGACTTCCGGAATTCGCGATGTCCAGCATCA 911  
QY 300 nGlnAlaProArgValGlyProAlaGlyGluHisArgLeuValThrValThrArgMetAs 320  
DB 912 TCAGGCGCCACGTGGGTGCGGAGGAGGAGCGCGGTGGTGTGCTGTACACAGATGGA 971  
QY 320 pGlyHisIleLeuGluGluAlaGluPheProGlyLeuProGluAlaLeuSerPheValAlaLe 340  
DB 972 CGGCCACATCTCTGGAAGCGGATTTCCGGGGCTGTGCTAGAGGCGCTCTTCTGGTGGCCCT 1031  
QY 340 uValAspGlyTyrPheArgLeuIleCysAspSerArgHisTyrPheCysLysGluValAl 360  
DB 1032 CGTGATGGGTACTTCCCGCTGATCTGGAGACTCCAGGCAATTTCTCCAAAGAGGTGGC 1091  
QY 360 aProProArgLeuLeuGluGluAlaAspValAlaCysHisGlyProIleThrLeuAspPh 380  
DB 1092 GCCGCCAAGGCTGCTGGAGAGGAGGAGCGGACTGTGTGCCATGAGCCATCACCTTGACTT 1151  
QY 380 eAlaIleHisLysLeuLysAlaAlaGlySerLeuProGlyThrTyrIleLeuArgArgSe 400  
DB 1152 TGCCATCCAGAGCTGAAGGCGGTGGCTCCCTCCAGGACCTATATTCTCGCGCGAG 1211  
QY 400 rProGlnAspTyrAspSerPheLeuLeuThrAlaCysValGlnThrProLeuGlyProAs 420  
DB 1212 CCGGAGAGACTATGACAGCTTCTTCTTAACCCCTGCGGCCAGACTCCCTGGGCCCGCA 1271  
QY 420 pTyrLysGlyCysLeuIleArgGlnAspProSerGlyAlaPheSerLeuValGlyLeuSe 440  
DB 1272 CTACAAAGGCTCCTCATCGCCAGAGACCCCAAGGGGCTTCTCTCCCTGGTGGCTCAG 1331  
QY 440 rGlnProHisArgSerLeuArgGlyIleuLeuAlaAlaCysTrpAsnSerGlyLeuArgVa 460  
DB 1332 CCAGGCCCAAGAGCTGCGGGAGCTGTTGCAGGCTCTGGATTTGGGCTGGAGAT 1391  
QY 460 lAspGlyAlaAlaLeuTyrLeuThrSerCysAlaProArgProLysGlnLysSerAs 480  
DB 1392 AGACGGTGTGCGCTGTACTTAACATCCGTGCGCTCCCAAGCCCAAGAAAGGCCAA 1451  
QY 480 ileuIleValIleArgArgGlyCysAsnProAlaProAlaProGlyCysSerProSerCy 500  
DB 1452 TTTGATCTGTGGCCAGAGGGGCTGCAACCCCGCGCTGCGCTGGCTCCCGGCTCG 1511  
QY 500 scYsaIaLeuThrGlnLeuSerPheHisThrIleProThrAspSerLeuGluTrpHisG 520  
DB 1512 CTGTGGCTGACACAGCTAGCTTCCACACAAATTCCAAAGGAGAGGCTGGAGTGGAGCA 1571
```

QY 520 uasnleuGlYhISglYserPheThrLysIlePheArgISerArgAGluValAs 540
 |||||
 Db 1572 GAACGTGGTACAGGCTCTTTTACCAAGATCTCGTGCGACAGCGGAGGCTGTGA 1631
 QY 540 pGlygluThrHisAspSerGluValLeuLeuLysValMetAspSerArgHisArgAsnCy 560
 |||||
 Db 1632 TGGGAGACATATGACTCGGAAGTCTCTCAAGGTCAATGACTCCAGACATCGGAACGTG 1691
 QY 560 smeGluserPheLeuGluValAlaSerLeuMetSerGluValSerTyrProHisLeuVal 580
 |||||
 Db 1692 CATGAGCTCTTTCTGGAGAGCCGCAAGCTTATGACCAATCTTACCGCACTGTGT 1751
 QY 580 lleuLeuHisISglYalCysMetAlaGlyAspSerIleMetValGlnGluPheValTyrIle 600
 |||||
 Db 1752 GTTACTGCACGCGCTGCTCATGGCTGGAGACAGCATGTGTGCAAGAAATTTGTGTACT 1811
 QY 600 uGlyAlaIleAspMetTyrLeuArgLysArgGlyHisLeuValSerAlaSerTyrPlySle 620
 |||||
 Db 1812 AGGAGCAATTTGATGATGACCGCAACCGTGGCCACTGCTGTGACGCACTGGAAACT 1871
 QY 620 uGlnValThrLysGlnLeuAlaTyrAlaLeuAsnTyrIleGlnLysPlySglYLeuProHis 640
 |||||
 Db 1872 GCAGGTGACCAAGACAGCTGGCATATGCCCTTAACCTTGGAGAGCAAAAGGCTTCTCA 1931
 QY 640 sGlyAsnValSerAlaArgLysValLeuLeuAlaArgGlnGlyLysPlyAsnProPr 660
 |||||
 Db 1932 CGGCAAGCTCTCACAGCGGAAGGTGCTGCTGCTGAGAGGGGTGATGGAAATCCACC 1991
 QY 660 oPheIleLysLeuSerAspProGlyValSerProThrValLeuSerLeuGluMetLeuTh 680
 |||||
 Db 1992 TTTATTTAAGCTGAGTGTCTGTGTGTGATCCACAGTGTGTGAGCCCTGGAAATGCTCAC 2051
 QY 680 rAspArgIleProThrPylAlaProGluCysLeuGlnGluAlaGlnThrLeuCysLeuG 700
 |||||
 Db 2052 CGACAGATACCTGGGGGGCCCGCAATGCTCCAGAGAGCTCAGACACTGCTTGA 2111
 QY 700 uAlaAspLysTrpGlyPheGlyAlaThrThrTrpGluValPheGlnArgGlyProAlaHis 720
 |||||
 Db 2112 GGGGCAAGAGGGGGCTTTGGAGCCACCACTGGAGGTGTCCCAAGGGGACCCGGCCA 2171
 QY 720 sIleThrSerLeuGlnProAlaLysLysLeuLysPheTyrGlnLysPlySglYLeuPro 740
 |||||
 Db 2172 CATACCTCGCTGGAGGCCCGCAAAAGCTGAACTTATGAGAGCAAGGAGACAGCTGCC 2231
 QY 740 oAlaLeuLysTrpThrGlnLeuAlaGlyLeuIleThrGlnCysMetAlaTyrAspProG 760
 |||||
 Db 2232 CGCTCTCAATGACAGACAGCTGGGAGCTTATCACAAGAGCATGGCGTATGATCTGG 2291
 QY 760 yArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIleThrSerAspTy 780
 |||||
 Db 2292 CCGGCGCCCTCTTCCGAGCTATCTCAGAGACCTCAACGCGCTCATTTACATCAGATT 2351
 QY 780 rGluLeuLeuSerAspProThrProGlyIleProSerProArgAspLysLeuCysValAl 800
 |||||
 Db 2352 CGAGCTCTCTCAGAGCCCAACACTGGCATCCGAGTCTTCGAAATAGAGCTGTCCGTA 2411
 QY 800 acGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGlnGluArgHisLeuLysTy 820
 |||||
 Db 2412 TGGGCGCCACTATCTCCCTGCCAGGACCCCGCATTTTCAGAGAGAGACACCTTAAGTA 2471
 QY 820 rIleSerLeuLeuGlnLysGlnAspPheGlySerValGlnLeuCysArgTyrAspProLe 840
 |||||
 Db 2472 CATCTCTTCTGGGCAAGGCAACTTTGGAGGCTGGAGCTGGCGCTATGACCCCT 2531
 QY 840 uGlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGlnHisSerValProAspG 860
 |||||
 Db 2532 GGGGAGACAAATACGGGAGCCCTGGTGGCAGTGAAGACACTAGACACGCGTCCAGACA 2591
 QY 860 nGlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuHisSerAspPheIleVal 880
 |||||
 Db 2592 GCAGAGGAGACTTCAGCGGAGATTCAGATCTTAAGGCTCTGACACAGCACTTCATGCT 2651
 QY 880 lLysTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuArgLeuValMetGluTy 900

Db 2652 CAAGTACCGGGAGTACACTATGGCCAGTCCAGAGCTGGCTTTGGATGAGATATA 2711
 QY 900 rLeuProSerGlyCysLeuArgAspLeuLeuGlnArgHisArgGlyLeuHisThrAspAr 920
 |||||
 Db 2712 CTTCGCCAGCGGTGCCCTCGCAGACTTCTCAGCGCATTCGCGGCGTCCAGACCGACCG 2771
 QY 920 gLeuLeuLeuPheAlaTrpGlnIleCysLysGlyMetGluTyrLeuGlnGlyAlaArgArgCy 940
 |||||
 Db 2772 CTTACTGCTGTGCTTGGCAGATCTGCAGAGGATGAGATGACTGGGTGGCCCGCTG 2831
 QY 940 sValHisArgAspLeuAlaAlaArgAsnIleLeuValGlnSerGluAlaHisValLysI 960
 |||||
 Db 2832 CGTACACCGTACCTCGGTGGTGGCCCAACATCTTGTGTGAGAGCGAGGCTCATGTGAGAT 2891
 QY 960 eAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAspTyrTyrValValArgG 980
 |||||
 Db 2892 CGCGGACTTCGCGCTCCCTAGAGTGTGCTGCCCTGGAAAGACTACTACGTGTCCCGCA 2951
 QY 980 uProGlyGlnSerProIlePheTyrPylAlaProGluSerLeuSerAspAsnIlePheSe 1000
 |||||
 Db 2952 GCTTGCCCAAGCCCATCTTTTGTGTATGCCCGGAGTCCCTATCTGACACATCTTCTC 3011
 QY 1000 rArgGlnSerAspValTrpSerPheGlyValValLeuTyrGlnLeuPheThrTyrCysAs 1020
 |||||
 Db 3012 CGGCCAATCTGACGTGTGAGACTTCGAGAGTGTGTGACAGACTTTCACCTACTGCGA 3071
 QY 1020 pLysSerCysSerProSerAlaGlnPheLeuArgMetMetGlyProGluArgGluGlyPr 1040
 |||||
 Db 3072 CAAGAGCTGACGCCCATCTCCGTGAGTCTCCGCGATATGAGGGCTGAGCGTGAAGAGAC 3131
 QY 1040 oProLeuCysArgLeuGlnLeuLeuAlaGlnGlyArgArgLeuProPropProTh 1060
 |||||
 Db 3132 CCGCTCTGCGCGCTCTCTGGAGCTGTGTGGAGAGAGGCGGAGCGCTCCACACTCCAC 3191
 QY 1060 rCysProThrGlnValGlnGlnLeuMetGlnLeuCysTrpAlaProGluProHisAspAr 1080
 |||||
 Db 3192 CTGGCCACCGAGGTTGAGGAGCTCATGACAGCTGTGTGGGCGCCAGCGCAGCAGCG 3251
 QY 1080 gProAlaPheAlaThrLeuSerProGlnLeuAspProLeuTrpArgGlyArgProGly 1099
 |||||
 Db 3252 GCCAGCTTGGCGACCTGTGAGCCCGCAGCTGGACCCGCTGTGGGTGAAGACCCGGA 3309
 RESULT 2
 MUSJAK3A 3528 bp mRNA Linear ROD 24-APR-1996
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Mus musculus (strain BALB/c x 129 F2) (clone: clone=s12-18) neonate
 thymus cDNA to mRNA.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathu; Muridae; Murinae; Mus.
 REFERENCE
 1 (bases 1 to 3528)
 AUTHORS
 Gurniak,C.B. and Berg,L.J.
 TITLE
 Murine JAK3 is preferentially expressed in hematopoietic tissues
 and lymphocyte precursor cells
 JOURNAL
 Blood 87 (8), 3151-3160 (1996)
 MEDLINE
 96184772
 PubMed
 8605329
 FEATURES
 location/Qualifiers
 1..3528
 /organism="Mus musculus"
 /strain="BALB/c x 129 F2"
 /db_xref="taxon:10090"
 /clone="clone=s12-18"
 /cell_type="T cell precursor"
 /tissue_type="thymus"
 /dev_stage="neonate"
 1..3528
 /gene="JAK3"
 gene

5'UTR
1. 71
/gene="JAK3"
/gene="JAK3"
72. .3374
/gene="JAK3"
/codon_start=1
/protein_id="AAC42085.1"
/db_xref="GI:1019910"
/translation="MAPSEETPLIPORCSLSSEACALHLLPPRGPPORRISFES
FGDYLAEDLCVRKACGILIPVHSLEFLATEDESCMPSPHICIEEDVDQVLYRL
REYPMDEGLETCHRFGRKDTLSIDLHLEHLEFADHRSGLRPLVGLSKMEOG
EFLSLAVIDLADMADEORARPEGLIKTYSTACILPSLRDIOGONFTRIRRTIV
LALRYVAQCDNRALMAKYLIDLERLHPATTEFRFRGLGAGEEPELLEACDNGI
SWSSGDDLEFQTECDPEIDVSTIKQARVPAGHRIIVTMDGHLEAFEGPLPE
ALSFVALVDGYFRILCDNRHYCKEVAPRLLIEAELECHPILDLFAIHKLAGSL
PGTYILRRSPDDYDSFLITACVQTPPLGPDYVGCILRODPSAFSLVGLSDQPSRLRL
LAACMSGLRVDAALNLTSCAPRPKSNLIYVRRCCTPAPAPGSCPCALTOIS
FRTIPDSELEHENGHSFPIKFRGRREYVDGETHSEVLKVMDSRHRNKMSFEL
EASLMSQVSPHLYLLHGNMAGDSIMVQFVYLGATDMYLRKRGHVYSKMLQVY
KQLYALNTLEDKGLPHGNVSARKYLLAREGGDGNPITKLSDRGVSPVLSLEMLTD
RIPWAPRCEQEAQTLCEADKMGATWVEFSGPAHITISLEPAKLKRYEEOGL
PALMTELAGLITQCMAYDGRSRFSFALIRDLNGLITSDVELSDPGLIPSPDEL
CGAQLYACODPAIFEERHLKYISLKGNGSYELCYDPLGDMTGPGLVAVKLOHS
GPDORDFOREIOLIKALHSDPIVYKRGVSGPGRSRLVMEYTPSGCLDPLORHR
ARLHDDLLLEFAMQICMEYLGARCYRDLARNTLISEAHYKINDDECLATLPL
GKDYVYAREPCOSPIFWYAPESLSDNITRSQSDVASEVVIYELFTYTDKSCSPSAE
LRMMGPEREGRPLCRLELLAEGRRLLPPPTCPTEVQELMOLCWAAPSPDHPAREGTLIS
POLDALMRGRPG"

3'UTR
3375. .3528
/gene="JAK3"
3517. .3522
/gene="JAK3"

polyA_signal
BASE COUNT 648 a 1104 c 1036 g 740 t
ORIGIN

Alignment Scores:
Pred. No.: 7.67e-205 Length: 3528
Score: 221.00 Matches: 1060
Percent Similarity: 96.89% Conservative: 0
Best Local Similarity: 96.89% Mismatches: 15
Query Match: 20.11% Indels: 34
DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x MUSTAK3A (1-3528)

QY 1 MetAlaProSerGluGluThrProLeuIleProGlnIleArgSerCysSerLeuSerSer 20
DB 72 ATGGGACCTCCCAAGTAGAGAGACACTGTGATCCCTGACGCTCTTGACGCTCTCATCC 131

QY 21 SerGluAlaGlyAlaLeuHisValLeuLeuProProAlaGlyProGlyProProGlnArg 40
DB 132 TCAGAGGCGAGGAGCCCTGTGATGTCTCTCCCGGGGACCTGGGCTCCCGCAGCGA 191

QY 41 LeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysValAlaArgAlaAlaVala 60
DB 192 TTGTCTATTTCTCTTTGGGAGCTACTTGGCTGAGGATTTATGTGTCTGCGAGCTGCCAAGGCC 251

QY 61 CysGlyIleLeuProValTyrHisSerLeuPheAlaLeuAlaThrGlnAspPheSerCys 80
DB 252 TGTGGGATCTGCTGCTTTATCATTCGCTTTTCGCTCTGCGACATGAGGACCTTCTCTTC 311

QY 81 TrpPheProProSerHisIlePheCysIleGluAspValAspThrGlnValLeuValTyr 100
DB 312 TGGTTTCCCCCAAGCACATCTTTCGATVAGAGAGAGCTGGACACTCACTCTTGGCTCTAC 371

QY 101 ArgLeuArgPheTyrPheProAspTrpPheGlyLeuGluThrCysHisArgPheGlyLeu 120
DB 372 AGGCTACGCTTTTATTTTCCCTACATGGTTTGGGCTGGAGACATGTACCGCTTTGGGCTG 431

QY 121 ArgGlyAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLeuPheAlaGln 140
DB 432 CCCAAGAAGATTGACAGGAGTGCATCTTACATGATGTTTAAACATCATCTTTGGCTGAG 491

QY 141 HisArgSerAspLeuValSerIleArgLeuProValGlyLeuSerMetGluGlnGly 160

|||||
DB 492 CACCCAGTACAGCTGTGAGTGGGCGCTCCGCGTGGGCTTACGATGAGAGAGAGCGA 551

QY 161 GluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGlnArgGlnArg 180
DB 552 GAGTTCTCTGAGCTGTGCGCTGCTGAGCTTGGCCAGATGGCTGTGAGAGAGCCCGCAGCC 611

QY 181 ProGlyGluLeuLeuHisThrValSerTyrLysAlaCysLeuProProSerLeuArgAsp 200
DB 612 CCAGGAGAGCTCTGTGAAGAGCGTCAAGTTACAAAGCTGTCTGCGCGCAGCTGGCGCAT 671

QY 201 ValIleGlnGlyGlnAspPheValThrArgGlnGlnIleArgArgThrValValLeuAla 220
DB 672 GTGATCCAGGGCCAAACTCTGTGACACCGCAGCGCGATCCGAGAGCCGTGGCTTTGGGG 731

QY 221 LeuLeuProCysGlyArgLeuProGlyArgPro-TyrAlaLeuMetAlaLysTyrIleLe 240
DB 732 CTGGC-CCGTGTGTGTGCTGCTGCGCAGGCGCAGCGCTCATGCGCATATATTTCT 790

QY 240 uAspLeuGluArgLeuHisProAlaAlaThrThrGlnThrPheArgValGlyLeuProG 260
DB 791 GGACCTGGAGCGGCTATCATCCAGCGGACACCGAGACCTTCGTTGGGCTCCCGGG 850

QY 260 ValAlaGlnGluArgProGlyLeuLeuArgValAlaGlyAspAsnGlyIlePro-TrpSerS 280
DB 851 CCGCCAGGAGGAGCGGGGCTTCTGCTGTGGCGGGGCAACGGCATCTC-CTGGAGCT 909

QY 280 eFAsn-Asp---GluLeuPheGlnThrPheCysAspPheProGluIleValAspValSer 298
DB 910 CCGG-GGACCGAGAGCTTTTCCAGACCTTCTGTGATTTCCGGMAATCTGATGTGCACG 968

QY 299 IleAsn-GlnAlaProArgValAlaGlyProAlaGlyGlnHisArgLeuValThrValThrAr 318
DB 969 ATCAA-GCAGGCCCAACGTTGGTGGTCCGCGAGGAGACCGCGCTGCTCATGTCCACG 1027

QY 318 GMetAspGlyHisIleLeuGluAlaGluPheProGlyLeuProGluAlaLeuSerPheVal 338
DB 1028 GATGAGCGGCGCATCTCGGAAGCGAGTTCCGGGGCGCTGAGGCGCTCTCTTTCGT 1087

QY 338 lAlaLeuValAspGlyTyrPheArgLeuIleCysAspSerArgHisTyrPheCysLysG 358
DB 1088 GGGCCCTGCGATGGGTACTTCCGCTGTGATCTGCGACCTCAGAGCATATTTCTGCAGAGA 1147

QY 358 uValAlaProProArgLeuLeuGluGluGluAlaAspVal--CysHisGlyProIleThr 377
DB 1148 GGTGGCGCGCCACGAGCTCTGTGAGAGAGAGCGGAG-GC-TGTCCATGTGACCAATCAGC 1205

QY 378 LeuAspPheAlaIleHisTyrLeuLysAlaAlaGlySerLeuProGlyThrTyrIleLeu 397
DB 1206 TTAGACTTTGCCATCCACAAGCTGAAGGCGGCTGCTCCCTCCAGGACCTATATTTCTC 1265

QY 398 ArgArgSerProGlnAspTyrAspSerPheLeuLeuThrAlaCysValGlnThrProLeu 417
DB 1266 CGCCGAGCGCGCAGAGACATAGACAGCTTCTTCTTACGCGCTCGCTCCAGACTCTCTTT 1325

QY 418 GlyProAspTyrLysGlyCysLeuIleArgGlnAspProSerGlyValaPheSerLeuVal 437
DB 1326 GCGCCCGAGTACAAAGGCGCTGCTCATCCGCGAGAGACCCAGGGGCGCTTCTCCCGGT 1385

QY 438 GlyLeuSerGlnProHisArgSerLeuArgGluLeuLeuAlaAlaCysTrpAsnSerGly 457
DB 1386 GGCCTCAGCCAGCCCGCACGAAGCCTGCGGAGCTCTTGCAGCGCTGGAATTTCTGGG 1445

QY 458 LeuArgValAspGlyAlaAlaLeuTyr-LeuThrSerCysCysAlaProArgProLysG 477
DB 1446 CTGCGAGTAGAGGTGCTCCCTGAA-CTTAACTCTCTCTGCGCTCCAGACCCAGAGA 1504

QY 477 uLysSerAsnLeuIleValValaArgArgLysAsn-ProAlaProAlaProGlyCysS 497
DB 1505 AAAGTTCAAATTGATCTGTGTGCGAAGGGGCTGAC-CCCGCGGCTGCGCCCTGGCTGCT 1563

QY 497 eTrProSerCysAsnAlaLeuThrGlnLeuSerPheHisThrIleProThrAspSerLeuG 517
|||||

Db 1564 CCCCCTCTGCTGTCGCTGACACAGCTGAGCTTCCACACATTCACAGGACAGCTGG 1623
 QY 517 TTTTPhisGluAsnLeuGlyHisGlySerPheThrLysIlePheArgGlySer-ArgArg 536
 Db 1624 AGTGGCAGAGAACCTGGGTGTCAGCGTTCTTTTACCAGATCTTCCTGGCGG-CAGGGGG 1682
 QY 537 GtValValAspGlyGlyThrHisAspSerGluValLeuLeuLysValMetAspSerArg 556
 Db 1683 GAGGTCCGTGATGGTGGAGACACATGACCTCGAAGTCCCTCGAAGGTATGACTCCGGA 1742
 QY 557 HisArgAsnGlyMetGlySer-PheLeuGluValAlaSerLeuMetSerGlnValSerTyr 576
 Db 1743 CATGGAACTGACATGAGATCTTTCTGGAAAGCCCAAGCTTGATGAGCCAGATATCTTAC 1802
 QY 577 ProHisLeuValLeuLeuHisGlyValCysMetAlaGlyAspSerIleMetValGlnGlu 596
 Db 1803 CCGACCTGGGTGTTACTGTCACAGGCGTGTGATGCTGGAGACATCATGATGGAGAA 1862
 QY 597 PheValTyrLeuGlyAlaAlaLeuAspMetTyrLeuArgGlyHisLeuValSerAla 616
 Db 1863 TTTGTGATCTAGAGACAAATTGACATGACTGGCAAGCTGGCCACCTGCTGCACACC 1922
 QY 617 SerTyrPylsLeuGlnValThrLysGlnLeuAlaTyrAlaLeuAsnTyrLeuGlnAspLys 636
 Db 1923 AGCTGGAACAGTCAAGGTGACCAAGCAGCTGCATATGCCCTTAACCTAGTGGAGACAAA 1982
 QY 637 GtLeuProHisGlyAsnValSerAlaArgLysValLeuLeuAlaArgGlnGlyAsp 656
 Db 1983 GGCCTTCTCAGCGCACAGCTGTCAGCAGGAGAGTGCCTCGCTCGTGGAGGGGGTGTAT 2042
 QY 657 GtLysAsnProPheIleLysLeuSerAspProGlyValSerProThrValLeuSerLeu 676
 Db 2043 GGGAAATCACCCTTTCATTAACCTAGTATGATCTGATGCTGATCCACAGTCTGAGCCTG 2102
 QY 677 GtLysLeuThrAspArgGlyLeuProTyrValAlaProGlyCysLeuGlnGluAlaGlnThr 696
 Db 2103 GAAATGCTCACCGACAGAAATACCTGGGTGGCCCCGGAATGCTTCACAGAGGCTCAGACA 2162
 QY 697 LeuCysLeuGlnAlaAspLysTyrGlyPheGlyAlaThrThrPheGlyValAlaPheGlnArg 716
 Db 2163 CTCGCTTGGAGGCTGACAAATGGGGCTTGGACCCACACAGTGGAGGTCTT-CAGGG 2221
 QY 717 -GlyProAlaHisIleThrSerLeuGluProAlaLysLysLeuLysPheTyrGluAspGly 736
 Db 2222 GGGACCCGCGCCACATCACTGCTGAGAGCCCGCAAAAAGCTGAAGTTCTATGAGACCA 2281
 QY 736 nglyGluLeuProAlaLeuLysTyrPheGlnLeuAlaGlyLeuIleThrGlnCysMetAla 756
 Db 2282 GGGACAGCTGCGCTCTCAATGAGACAGAACTGGCGGACTTATCACACAGTGCATGCG 2341
 QY 756 atYAspProGlyArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIle 776
 Db 2342 GTATGATCTGCGCGCGCGCGCTCTTCGAGCTATCTCTCAGAACTCAACGGGCTCAT 2401
 QY 776 erHisSerAspTyrGluLeuLeuSerAspProThrProGlyIleProSerProArgAspGly 796
 Db 2402 TACATCAGATTACAGAGCTCTCTCAGACCCACACACCTGGCATCCCGAGTCTCGAGATGA 2461
 QY 796 uLeuGlyValAlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGlnGluArg 816
 Db 2462 GCTCTGGG--GT-GGGCGCCGAGCTCTATGCTCTGCGAGAGCCCGCATATTCGAGGAGAG 2518
 QY 816 gHisLeuLysTyrIleSerLeuLeuGlyLysGlyAsnPheGlySerValGluLeuGlyAsp 836
 Db 2519 ACACCTTAAGTACATCTCTTGGTGGGCAAGGCAACTTGGCGAGCGGAGCTGTGGCG 2578
 QY 836 gTyrAspProLeuGlyLysAsnThrGlyProLeuValAlaValLysGlnLeuGlnHis 856
 Db 2579 CTATGAGCCCTCGGGGAGACATAGGAGCCCTGCTGTCAGTGAACAGCTACAGCACAG 2638
 QY 856 rVal-ProAspGlnGlnArgAspPheGlnArgGlnGlnIleGlnIleLeuLysAlaLeuHis 876
 Db 2639 CGG-GCCAGACACAGAGGAGCTTCCAGCGGAGATTCAAGATCTTAAGGCTCTGCACA 2697

QY 876 erAspPheIleValLysTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuArgL 896
 Db 2698 GCGACTTTCATGTCAAATGACGCGGAGATGACCTATGAGGCGCAGGTGCGACAGCTGCGGT 2757
 QY 896 euValMetGlnTyrLeuProSerGlyCysLeuArgAspLeu-LeuGlnArgHis----Ar 914
 Db 2758 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2816
 QY 914 gGlyLeuHisThrAspArgLeuLeuLeuPheAlaTyrGlnIleCysLysGlyMetGluTyr 934
 Db 2817 CGC-CTGCACACCGACCGCTCTACTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2875
 QY 934 rLeuGlyAlaArgArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGluLeu 954
 Db 2876 CTTGGGTGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2935
 QY 954 rGluAlaHisValLysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAs 974
 Db 2936 CGAGCTCATGTCAAAGATCGCGGAGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2995
 QY 974 pTyrTyrValValArgGluProGlyGlnSerProIlePheTyrPylalProGlnSerLe 994
 Db 2996 CTACTACCTGCTCGCGCGAGCTGCGCAAGCCCATCTTTGGATGCTGCTGCTGCTGCTGCTG 3055
 QY 994 uSerAspAsnIlePheSerArgLysSerAspValTyrPhePheGlyValValLeuTyrGly 1014
 Db 3056 ATCTGACACACATCTTCTCCGCCCAATGACGTGTGAGAGCTTGGAGAGTGTGTGACGA 3115
 QY 1014 uLeuPheThrTyrCysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetGly 1034
 Db 3116 GCTCTTCACTACCTGCGCAGACAGCTGACGCCCATCCGCTGAGTTCCTGCGCAGATGATGG 3175
 QY 1034 yProGluArgGlnGlyProProLeuGlyArgLeuLeuGlnIleLeuMetGlnLeuCysTyrPyl 1054
 Db 3176 GCTGAGGTGAGAGAGCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3235
 QY 1054 gLeuProProProProThrCysProThrGluValGlnGlnLeuMetGlnLeuCysTyrPyl 1074
 Db 3236 CCTCCACACACTCCACACTCCACACTCCACACTCCACACTCCACACTCCACACTCCACACTCC 3295
 QY 1074 apro 1075
 Db 3296 GCCC 3299
 RESULT 3
 MUSJAK3H
 LOCUS
 DEFINITION Mus musculus (clone 32D5-1). protein tyrosine kinase (JAK3) mRNA.
 ACCESSION L33768
 VERSION L33768.1 GI:508533
 KEYWORDS protein tyrosine kinase.
 SOURCE Mus musculus cDNA to mRNA.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 AUTHORS Rane,S.G. and Reddy,E.P.
 TITLE JAK3: a novel JAK kinase associated with terminal differentiation
 of hematopoietic cells
 JOURNAL Oncogene 9 (8), 2415-2423 (1994)
 MEDLINE 9430920
 PUBMED 7518579
 FEATURES
 source
 location/Qualifiers
 I..4016
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="32D5-1"
 /cell_line="32Dcl3"
 /cissue_lib="lambda gtl1"
 7..3906
 /note="protein tyrosine kinase"
 /codon_start=1

```

/product="JAK3"
/protein_id="AA21415.1"
/db_xref="GI:537607"
/translation="MAPSEEPRLIQRSCSSSSAGALHVLPPRGPPQRLSFS
FGDYLABEDCVRAKACALIPVHSLFALATEDSCWFPSPHLEEDVDYVRL
RYFPMFGLETCRHRGLRDLRLDHLVLEHLHQAHRSDIYSGRLPVGLSMKQ
EFLISLAVLDLAQARQAPRGLLKTVAQKALPSLIDVQGNFVRRIRRRQ
LALRWSPARPTAHQAVYDLERLHPATETTFVGLPQOEEFGLRVAGDNGIS
MSSGQDEVLGILRSVGRVAMRLVKPVRTIRKVASPCACCTGGGGLKQNG
MELPQGLTWGVTBRVRLDRGRGTSGNDPPTKSPVSPRLSPSPKPKMRBG
GRRGCGASRSVPMILSLFPPFNGPRQGFSSSGCPGTHEVAPMPKQKSA
CLSSAVTKGRVPLRLKCLSTFLQVTFCDPFLVDVSIKQPTGCGRHRRLV
DGHILAEPRGLEALSFVALVDGYRLICDSRHHFCKEVAAPRLIEBEALCH
LGSASGQOMGMAARTVGLMTLSMPRLDIAHKLKAASLPGLYILRSPQDVS
FLTACVQPLGPDYKGLIROPDSGAFSLVASASTEGSCLOLEGLRVDGAL
NLTSACAPRREKSNLIYVRGCTPAPRAGSPSCALFOLSFHITPPDSCHEM
GSPFKIFGRRRREVVDGTHDEVLKLVMSDRHNRMEFLAASLSMGSYPIHVL
HGVCMAGDSIMQEPYLGALIMYLRKRLHLYSASKLOVTKQIALVLYLEDKL
GNVSAKVLILAREGGDGNPFTIKLSDPGSPVLSLEMLTDRIPMVAPECLOE
LEADKWFEGATTWEEVSGPAHITSLERPAKLFYEDQQLPALMTELAGLITC
YDGRPSFRALRLDLNGLITSDYELLQPHLSRVLKEMSCAMRALCLPGRHTE
HLKYLISLKGNGSVELCRDPLDNTGRLVAVKOHLGSPQOORPOREIOLK
SDIYKRGVSYGPGROSRLVMEYLPSGLRDLFOLHNAALHTRILLFNAOIT
EYLGARCVHRLAARLITVESEAHYKIDPCLRLALPLKRYTYVREPGSPIT
PRLSDNIFSRSDVMSRGVLYELFTTODKSCSFAELRLMRGERRGPRCLLEL
LABGRLLPRLPPTCEVDELQVAPSRITTGQPSAP"
misc_feature
2199..3024
       /note="JAK homology domain II (JH2)"
misc_feature
3076..3876
       /note="JAK homology domain I (JH1)"
BASE COUNT      742 a 1203 c 1235 g      836 t
ORIGIN

Alignment Scores:
Pred. No.:      7.77e-165      Length:      4016
Score:          180.00      Matches:      529
Percent Similarity: 97.24%      Conservative: 9
Best Local Similarity: 97.24%      Mismatches: 0
Query Match:    16.388      Indels:      15
DB:             10      Gaps:      0

US-09-397-967-16 (1-1099) x MUSJAK3H (1-4016)

QY 535 ARGARGLVALVALASPLGLUPTHNLSAASPERGLVALLEULEUUSYVALMETASP 554
    |||||
Db 2257 AGCGCGGAGGTGCTGATGCTGACACATGACTCGGAAGTCTCTCGAAGGTTCATGGAC 2316

QY 555 SERATGHNISATGASNCYSMETGLUSERPHEUGLUALAALSERLEUMETSERGINVAL 574
    |||||
Db 2317 TCCAGACATCGGAACATGATGAGTCTTTTGGAAAGCGGCAAGCTTGATGAGCCACAGA 2376

QY 575 SERTYPRONHISLEUVALLEULEUHNISGLYVALCYSMETALAGLYASPSERTILEMETVAL 594
    |||||
Db 2377 TCCGTACCCCGACCTGGTGTACTGACACGCGCTCTGATGGCTGGAGACATCATGGTG 2436

QY 595 GINGLUPHEVALTYRLEUGLYALALIASPMETTYRLEUATRGYLAARGLYHISLEUVAL 614
    |||||
Db 2437 CAGGAAATTGTGTATCTAGAGACATTCATGATCTGGCGCAACCGTGGCCACCTGGTGG 2496

QY 615 SERALASERTYRPLYSLEUGLINALTHIRLYSGINLEUALATYRVALALEUASNTRYRLEUGLI 634
    |||||
Db 2497 TCAGCCACGCTGGAACTGACGACCAAGAGCTGGCATATAGCCCTTAACACTTTGGAG 2556

QY 635 ASPLYSGLYLEUPRONHISGLYSVALSERALATGYLSVALLEULEUALAARGLYGLU 654
    |||||
Db 2557 GACCAAGGCTTCTTCACGCGCAACGCTTCACGACGGAAGGTCCTCTGCTGCTGAGGGG 2616

QY 655 GLYASPLGLYASNPPOPHETILEULYSLEUSERASPPROGLYVALISERPROTHYRLEU 674
    |||||
Db 2617 GGTATGGGAATCCACCTTTCTATTAAGCTGATGATCTCTGGGTGATGCCACTGTGCTG 2676

QY 675 SERLEUGLUMETLEUTHASPARGLTLEPROTPVALALAPROGLYCYLSLEUGLINALA 694
    |||||

```

```

Db 2677 AGCGTGAANAATGCTCAACCGACAGATACCTGGGTGGCCCCCGAATGTCTCCAGAGGCT 2736
QY 695 GLIPTHLEUCYSLEUGLUALAASPLYSYRPGLYPHEGLVALATHRTHTRPGLUVALPHE 714
    |||||
Db 2737 CACACACTGCTGCTGGAGGCTGACNAGTGGGGCTTTGGAGCCACACACGTTGGAGAGTGT - 2795

QY 715 GLNARG-GLYPROALAHISITLETHSERLEUGLUPROALALYSLYSLEULYPHETYRGL 734
    |||||
Db 2796 CACGGGGGACCCCGCCACATCACCTCGTGAGGCCGCCCAAAAGCTGAAGTTCTATCA 2855

QY 734 USAPGINGLYGINLEUPROALALEUUYSTRPHRGILUEUALAGLYLEUILLETHRGINCY 754
    |||||
Db 2856 GGACCGAGGACAGAGCTCCCGCTCTCAATGGACAGACTGGGGGAGCTTATACACAGAG 2915

QY 754 SHELATYTRASPPROGLYARGARGPROSERPHEARGALILEUATGASPLAENGL 774
    |||||
Db 2916 CATGGGCTATGATCTCGGCGGCGGCCCTCTCCAGAGTATCTCAGAGACTTCACAGG 2975

QY 774 YLEUILLETHRSERASPPYRGLILEULEUSERASPPROTHYRPROGLYILLEPROSERPROAR 794
    |||||
Db 2976 CCTCATTTACATGACATTTACGAGCTCTCCCA-GACCCACACCTGGCATCCGAGTCTCG 3034

QY 794 GASPGLULEUCYSVALALAGLYLAGLILEUTYRALCYSGINASPPROALALIEPHEGL 814
    |||||
Db 3035 AGATGAGCTGTGCG--TT--GGCGCCAGCTGATGCTCGCCAGACCCCGCATATTC-CA 3089

QY 814 UGLUARGHNISLEUUYSTRYILLESERLEUGLILYLSGLYASNPHGLYSERVALGLUL 834
    |||||
Db 3090 GGAGAGACACTTAAAGTACATCTCTTGTGGGCAAGGCAACTTGGACACGCTGAGACT 3149

QY 834 UCYSARGTYRASPPOLEUGLILYASPSNTHRCILYPROLEUVALALVALYLSGINLEUG 854
    |||||
Db 3150 GTGCCGCTATGACCCCC--TG-GACATATACGGAGACCTCGGTGGCAGTGAANAACATCA 3206

QY 854 HNISERVAL-PROASPLINGLARGASRPHENGLARGLIULEGINLLEULYSLAL 874
    |||||
Db 3207 GCACAGCGG-CCACACACGACGAGAGGACTTCCAGCGGAGATTCAGATCTTAAGGCTC 3265

QY 874 EUNHISERASRPHLEVALILYSTYRARGLYVALISERTYRGLYPROGLYARGINSERL 894
    |||||
Db 3266 TGACAGAGGACTTATCTGTCAGATGCCGGGAGTCAAGTATGGGCGAGGTGCCAGAGCC 3325

QY 894 EULATGLEUVALMETGLUTYRLEUPROSERGLYCYLSLEUARGSPLEU-LEUGLARGHIS 913
    |||||
Db 3326 TCGGCTTGATGATGAGTACCTGCCACAGCGCTGCGTCCGAGA-CTTCCTGACACGCCAT 3384

QY 914 --ARGGLY-LEUHNISTHASPARGLILEULEUHEALATRRGINLIECYLYSGLYMET 932
    |||||
Db 3385 CGCGGGGCTGACACACCGGACCGCTCTGCTGTTGCCATGGCAGATCTGCAGAGGCAATG 3444

QY 933 GLUTYRLEUGLYALATARGARGCYVALHISARGASPLEUUALAARGASNILLEUVAL 952
    |||||
Db 3445 GAGTAACTGGGTGGCGCGCGCTGCGGTACACCGTACCTGGCTGGCCACATCTTG6TG 3504

QY 953 GLUSERGLUALAHISVALYSILEALASRPHESGLYLEUALALYSLEULEUPROLEUGLY 972
    |||||
Db 3505 GAGAGCGAGGCTCATGTGAAGATCGCGGACTTGGCTCGGTAGCGTGGCCCTG6GA 3564

QY 973 LYSASPPYRTRYVALVALARGGLUPROGLYGINSEPROILEPHEPTYRVALAPROGLU 992
    |||||
Db 3565 AAGGACTACTACGTGGTCCGCGGACCGGCCAAAGCCCAATCTTTTGGTATGCCGCCGAG 3624

QY 993 SERLEUSERASPSNIIIEPHESEFARGLISERASRYVALTRPSERPHEGLYVALVALLEU 1012
    |||||
Db 3625 TCCCTATCTGACACATCTTCTCCGCGCATCTACCTGTGAGACTTCCGAGTGG6TGTG 3684

QY 1013 TYRGLYUUPHEPTYRTRYCYASRPLYSERCYSERPROSERALAGLUPHELEUATGMEC 1032
    |||||
Db 3685 TACAGGCTTTCACCTACTCGACAAAGAGCTGCAGCCCATCCGCTGAGTTCTGGCGATG 3744

QY 1033 METGLYPROGLUARGGLUGLYPROPOLEUCYSARGLEULEUGLILEUVALAGLUGLY 1052
    |||||
Db 3745 ATGGGCGCTGACGCTGANAAGACCCCGCTCTGCGGCTCTGGAGCTGTGGGACAGAGGC 3804

```


	RESULT 4	
	LOCUS RATAJK3	
	DEFINITION Rat mRNA for protein-tyrosine kinase, complete cds.	
	ACCESSION D28508	
	VERSION D28508.1 GI:485811	
	KEYWORDS JAK family; protein-tyrosine kinase.	
	SOURCE Rattus norvegicus adult spleen cDNA to mRNA, clone JAK3.	
	ORGANISM Rattus norvegicus	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;	
	Rattus.	
	1 (sites)	
REFERENCE	Takahashi,T. and Shirasawa,T.	
AUTHORS	Molecular cloning of rat JAK3, a novel member of the JAK family of	
TITLE	protein tyrosine kinases	
JOURNAL	FES Lett. 342 (2), 124-128 (1994)	
MEDLINE	94192816	
REFERENCE	2 (bases 1 to 3778)	
AUTHORS	Shirasawa,T.	
JOURNAL	Submitted	
COMMENT	Submitted (10-Feb-1994) to DDBJ by:	
	Takuiji Shirasawa	
	Department of Molecular Pathology	
	Tokyo Metropolitan Institute of Gerontology	
	35-2 Sakae-cho, Itabashi-ku	
	Tokyo 173	
	Japan	
	Phone: 03-3964-3241 x3034	
	Fax: 03-3579-4776.	
FEATURES	Location/Organism	
source	1..3778	
	/organism="Rattus norvegicus"	
	/db_xref="taxon:10116"	
	/tissue_type="spleen"	
	/dev_stage="adult"	
	1..3778	
	/gene="JAK3"	
	61..3363	
	/gene="JAK3"	
	/codon_start=1	
	/evidence="experimental	
	/product="protein-tyrosine kinase"	
	/protein_id="BAA05868.1"	
	/db_xref="GI:485812"	
	/translation="MAPSPETPLISORCSLSSEAGALHVLPPRGPGPQRLSFPS	
	/EGDYLADICVRAKAGGILPVYHSIFLALATEDSCMPRESIISIDVDQIVLYRI	
	RPYPGRGTGETCHRGFGLHKDLTSAIIDVHLHPNHRSDIASIGRLPYGLSLDKDG	
	EPFLSIADVLDLMARKQAKQAPRELKTSYKACIPSLRIDIOSQSVTRIRIRIVV	
	QALPAVSSLPSPRALAKAYTLDELRRPAPATFSFLGLDGADEEGCLRYSGDNLI	
	AMSSADELFQFCDFEIVDVYSIKQARVPAGEHRIVLTITRDGHILEAFGLPE	
	ALSFVAIDVGFRFLICDSRHFECKEVAPRLLEDAELCNCHPIITLDEAIHLKLAAGS	
	PSTYILRSPODYOSFLTACTVPTLGFDGYGCLIRDPSCAFSLVIGLSOHLRSLOE	
	LTCMCHSGLOYDALNTSCCPVPRKKSMLIVRRNRNPAPAGVSPSCALTKLSL	
	FHTIPADSLFMHNELNGHSPTKFIHGHRVGVDDETHTEVELLVKMSRHONCESP	
	EAASIMSOVSYPHYLLHGVCMAGDSIMVOEFYVIGAITDYILKRKHLYPSMKLOYTV	
	KOLAVAINYLEDKGILPHGNVSARKVLLAREGVDSNPPTKLSIDGVSPTVISLEMTTD	
	RIPWAPAELEGIAQMAYDPMGRSFRAILRDNLGITSDYELSDIPKPKFYEDRGQL	
	PAKTKTEGLELAQCMAYDPMGRSFRAILRDNLGITSDYELSDIPKPKFYEDRGQL	
	GCGAIVACOPALPFEEHRLKYISLLGGNGFVGLCYDPLGNTGPGLVAVKOLHSHS	
	GPGRODRPREFRQITLKALHCPDIVKRVGVSGPGKQELRLVMETLPSGCLDFLRH	
	ARIHLNDRLILRAMQICKRMETIGARCRCHROLARNRTIVSEAHVKITADIELATLP	
	KRDYIVVRPGQSPLFWAPESLSDNIITSKSDVWSBEVYLIELFTYSDDKSCSTEPE	
	LRMGPERGSPICHLLELLALEGRRLPPPTPCFTEVQELMQLCWSPNQDPAPDITLS	
	FOLDLMRGSPg"	
BASE COUNT	729 a 1172 c 1060 g 817 t	
ORIGIN	*	

Alignment Scores:		9.01e-53	Length: 3778
Pred. No.:	65.00	Matches:	99
Score:	98.02%	Conservative:	0
Percent Similarity:	98.02%	Mismatches:	1
Best Local Similarity:	98.02%	Indels:	2
Query Match:	5.91%	Gaps:	0
DB:	10		
US-09-397-967-16 (1-1099) x RATJAK3 (1-3778).			
QY	919	AspargLeuLeuPheAlaTrpGlnIleCysIysGlyMetGluTyrLeuGlyAlaArg	938
DB	2818	GACCCCTCAAGCTCTTGGCCCTGGAGATCTGCACAGGCGCATGGAAGTACCTGGCGCACGC	2877
QY	939	ArgCysValIHisArgAspLeuAlaAlaIArgAsnIleLeuValGluSerGluAlaHisVal	958
DB	2878	CCCTCGTGTCATCTGACCTGGCTGGCGCAACATCTGTGTGAGAGATGAGGCGCATGTG	2937
QY	959	LysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAspTyrTyrValVal	978
DB	2938	AAGATTGCTGACTTGGCGCTCGCTAAGCTGGCCCTCTTGGAAAGACTACTACTGGTC	2997
QY	979	ArgGluProGlyIysIleSerProIlePheTrpTyrIleAlaProGluSerLeuSerAspAsnI	998
DB	2998	CGCGT-GCCGGGSCCAAGTCCCATCTTCTGGTACGCCCTTGATGCCATATCGACACAT	3056
QY	998	ePheSerArgGlnSerAspValTyrPhePheGlyValValLeuTyrGluLeuPheThrTy	1018
DB	3057	CTTCTGCCCAATCTGACGTGTGAGAGCTTGGCGGTGGCTGTATGACTCTTCACCTA	3116
QY	1018	r 1018	
DB	3117	C 3117	
RESULT 5			
LOCUS	AX203087	778 bp	DNA linear PAT 30-AUG-2001
DEFINITION	Sequence 8 from Patent WO0152892.		
ACCESSION	AX203087		
VERSION	AX203087.1	GI:15392445	
KEYWORDS			
SOURCE	human.		
ORGANISM	human sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
TITLE	1 (bases 1 to 778)		
JOURNAL	Jak/stat pathway inhibitors and the uses thereof		
	Patent: WO 0152892-A 8 26-JUL-2001;		
	GENZYME CORPORATION (US)		
FEATURES	Location/Qualifiers		
source	1..778		
BASE COUNT	131 a 265 c 219 g 163 t		
ORIGIN	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
Alignment Scores:			
Pred. No.:	2.14e-50	Length:	778
Score:	62.00	Matches:	96
Percent Similarity:	97.96%	Conservative:	0
Best Local Similarity:	97.96%	Mismatches:	2
Query Match:	5.64%	Indels:	2
DB:	6	Gaps:	0
US-09-397-967-16 (1-1099) x AX203087 (1-778)			
QY	938	ArgArgCysValIHisArgAspLeuAlaAlaIArgAsnIleLeuValGluSerGluAlaHis	957
DB	103	CGCCGCTGCTGTCACCGCGACCTGGCGCCCGAACAATCTCTCGGAGAGACGAGGCAC	162
QY	958	ValIysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAspTyrTyrVal	977

[illegible]

VERSION	KEYWORDS	UT31601.1	GI:1039418
SOURCE	ORGANISM	Homo sapiens.	
REFERENCE	AUTHORS	Liu, E.T., A kinase-deficient splice variant of the human JAK3 is expressed in hematopoietic and epithelial cancer cells J. Biol. Chem. 270 (42), 25028-25036 (1995)	
TITLE		2 (sites)	
JOURNAL		Wittuhn, B.A., Silvennoinen, O., Mura, O., Lai, K.S., Cwik, C., Liu, E.T. and Ihle, J.N.	
MEDLINE		Involvement of the Jak-3 Janus kinase in signalling by interleukins 2 and 4 in lymphoid and myeloid cells	
PUBMED		Nature 370 (6485), 153-157 (1994)	
REFERENCE		8022486	
AUTHORS		3 (bases 1 to 3620)	
TITLE		Liu, E.T.	
JOURNAL		Direct Submision	
MEDLINE		Submitted (14-JUN-1995) Edison T. Liu, Lindeberger Comprehensive	
PUBMED		Cancer Center, University of North Carolina, Chapel Hill, NC	
REFERENCE		27599-7295 USA	
AUTHORS		Wittuhn, B.A. Nature 370, 153-157, 1994.	
TITLE		Location/Qualifiers	
JOURNAL		1..3620	
MEDLINE		/organism="Homo sapiens"	
PUBMED		/db_xref="taxon:9606"	
REFERENCE		/chromosome="19"	
AUTHORS		/map="19p12"	
TITLE		/clone="SKR-3"	
JOURNAL		/cell_type="SKBR-3 breast cancer cells"	
MEDLINE		1..3620	
PUBMED		/gene="jak3b"	
REFERENCE		96..3380	
AUTHORS		/function="tyrosine protein kinase"	
TITLE		/codon_start=1	
JOURNAL		/product="JAK3"	
MEDLINE		/protein_id="AAC50226.1"	
PUBMED		/db_xref="GI:1039418"	
REFERENCE		/translation="MAPPSEETPLIPRSCSLSTEGALHYLLPARGPQPRRLSFS FGDHLAEDICVOAKASGITLPVYHSALATPDLSCPPHSIFSVSDASTOVLIRI REYPMNGLCKRGLRDLASLIIDLPVLEHLFRQHSRDSVSRVGLSLKROG ECLSLAVIDLARMAREQRPGLKTVSYKACLPISLNDLIGLSFVRRIRIRVR RALRVAACQADRLSLAKYIMDLERIDPAGAEITHVGLPGLGHDGLGLRVAGD GGIATWGEDEVLQPCDPEIYDISTSQAPRVGAPAGEHRLVYTRTDNQLLEAFPG LPEALSVLALVDGFLRTDSQHFECEVAPPLLEBVAEGOCGPTTLDPAFNKRLS GSRGSLVLRSPDGFLLTVCVONQPDGDKGLIRSPGCTFLVGLSRPSSLT RELATQMDGGLHVDGAVVLTSCPIPREKSNLIVQKRGSPPTSLVQPOSOQL SOMPEHKLIPADLSLEHENTGHGSPKTIKRCRHEVVDGAKRTEVLLKVMADKHKNM ESPLEASLSMOVSRYRLVLLBVCVAGADSTWQERVHLGALIDYLRKGLHVPASMK LOYKQALAVLYLIEDGPHGNVSAKRYLLAEGADGSPFIKLSDPGSPAVLSLE MLDRIITWAPBECULREKQVTLSEADKMGATWEEFSGVTMPIKSLDPKALQFED RQQLPARKMETALLIQQCEWYEPVORPSFRAVIRLNLSSIDYELLDSPRGALAP RDGLMGNAQLYACODPTIFEEHRLKYSIOLQGNQSGSVFLCSDYELDQNTCALVAKO LHNSGPDQORDFOREIOLIKALSHDFIVRKCVSGPGOSIRLVMEYIIPSCLLVDPL ORRRARLDARLLSLKALSHDIFIVRKCVSGPGOSIRLVMEYIIPSCLLVDPL LPLDKRYVYVREPGOSPFIWYAPESLSNLISROSDWVSFVLYELFTVCDKSCSP SABTEIRMGCEERYPALCRLLELLEGGRLPAPACPAPEVSAAGLASVOSVDNAGVS GKPAQA"	
BASE COUNT		670 a 1152 c 1076 g 722 t	
ORIGIN			
Alignment Scores:		7.31e-50	Length: 3620
Score: 97.96%		62.00	Matches: 96
Percent Similarity:		97.96%	Conservative: 0

Best Local Similarity: 97.96% Mismatches: 1
 Query Match: 5.64% Indels: 2
 DB: 9 Gaps: 0

US-09-397-967-16 (1-1099) x HSU31601 (1-3620)

OY 938 ArgArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGluSerGluAlaHis 957
 |||||||
 DB 2922 CGCCGCTCGGACCGGACCTGGCCGCCGGAACATCCCTGCGAGAGGAGGACACAC 2981
 |||||||
 OY 958 ValIysIleAlaAspPheGlyLeuAlaLysLeuLeuPro-LeuGlyLysAspTyrTyrVa 977
 |||||||
 DB 2982 GTCAGATCGCTGAGCTTGGCTTACCTAGCTGCTGCGCTTGCAC-AAAGACTACTAGT 3040
 |||||||
 OY 977 lValArgGluProGlyGlnSerProIlePheTyrTyrAlaProGluSerLeuSerAspAs 997
 |||||||
 DB 3041 GGTCCGCGAGCCGAGCCAGAGCCCATTTTCTGATGCCGCCGATCCCTCGGACAA 3100
 |||||||
 OY 997 nIlePheSerArgGlnSerAspValTyrSerPheGlyValValLeuTyrGluLeuPheTh 1017
 |||||||
 DB 3101 CATCTCTCTCGCCAGCTGAGCTGTGAGAGCTTCGCGGTCTCTGTACGAGAGCTCTTAC 3160
 |||||||
 OY 1017 rTyrCysAspLysSerCysSerProSerAlaGluPheLeuArgMetGly 1034
 |||||||
 DB 3161 CTACTGCGACAAAGCTGACGCCCTCGGCGAGTTCTCGCGATGATGGGA 3212
 |||||||

RESULT 8
 AX203081 4064 bp DNA linear PAT 30-AUG-2001
 LOCUS Sequence 2 from Patent WO0152892.
 DEFINITION AX203081
 ACCESSION AX203081
 VERSION AX203081.1 GI:15392438
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 4064)
 AUTHORS Vasilos, G.
 TITLE Jak/stat pathway inhibitors and the uses thereof
 JOURNAL Patent: WO 0152892-A 2 26-JUL-2001;
 GENZYME CORPORATION (US)
 FEATURES
 source 1. 4064
 Location/Qualifiers
 BASE COUNT 746 a 1292 c 1168 g 858 t
 ORIGIN

Alignment Scores:
 Pred. No.: 8.02e-50 Length: 4064
 Score: 62.00 Matches: 96
 Percent Similarity: 97.96% Conservative: 0
 Best Local Similarity: 97.96% Mismatches: 1
 Query Match: 5.64% Indels: 2
 DB: 6 Gaps: 0

US-09-397-967-16 (1-1099) x AX203081 (1-4064)

OY 938 ArgArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGluSerGluAlaHis 957
 |||||||
 DB 2922 CGCCGCTCGGACCGGACCTGGCCGCCGGAACATCCCTGCGAGAGGAGGACAC 2981
 |||||||
 OY 958 ValIysIleAlaAspPheGlyLeuAlaLysLeuLeuPro-LeuGlyLysAspTyrTyrVa 977
 |||||||
 DB 2982 GTCAGATCGCTGAGCTTGGCTTACCTAGCTGCTGCGCTTGCAC-AAAGACTACTAGT 3040
 |||||||
 OY 977 lValArgGluProGlyGlnSerProIlePheTyrTyrAlaProGluSerLeuSerAspAs 997
 |||||||
 DB 3041 GGTCCGCGAGCCGAGCCAGAGCCCATTTTCTGATGCCGCCGATCCCTCGGACAA 3100
 |||||||
 OY 997 nIlePheSerArgGlnSerAspValTyrSerPheGlyValValLeuTyrGluLeuPheTh 1017
 |||||||

DB 3101 CATCTTCTCGCCAGCTGACAGCTGTGAGGCTTGGGCTCTCTGTACAGACTCTTTCAC 3160
 |||||||
 OY 1017 rTyrCysAspLysSerCysSerProSerAlaGluPheLeuArgMetGly 1034
 |||||||
 DB 3161 CTACTGCGACAAAGCTGACGCCCTCGGCGAGTTCTCGCGATGATGGGA 3212
 |||||||

RESULT 9
 HSU09607 4064 bp mRNA linear PRI 09-JUL-1994
 LOCUS Human JAK family protein tyrosine kinase (JAK3) mRNA, complete cds.
 DEFINITION Human JAK family protein tyrosine kinase (JAK3) mRNA, complete cds.
 ACCESSION U09607
 VERSION U09607.1 GI:508730
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1124)
 AUTHORS Kawamura, M., McVicar, D.W., Johnston, J.A., Blake, T.B., Chen, Y.,
 Lal, B.K., Lloyd, A.R., Kelvin, D.J., Staples, J.E., Ortaldo, J.R. and
 O'Shea, J.
 TITLE Molecular cloning of L-JAK, a Janus family protein-tyrosine kinase
 expressed in natural killer cells and activated leukocytes
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 91 (14), 6374-6378 (1994)
 MEDLINE 94294384
 PUBMED 8022790
 REFERENCE 2 (bases 1 to 4064)
 AUTHORS O'Shea, J.
 TITLE Direct Submission
 JOURNAL Submitted (11-MAY-1994) John O'Shea, Leukocyte Cell Biology Section,
 Fderick, MD 21702, USA
 FEATURES
 source 1. 4064
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /cell_line="YT, HUT-78"
 /cell_type="natural killer cell, PMA activated T cells"
 1. 4064
 /gene="JAK3"
 96.3470
 /gene="JAK3"
 /codon_start=1
 /product="JAK family protein tyrosine kinase"
 /protein_id="AA19626.1"
 /db_xref="GI:508731"
 /translation="MAPPSEPTLIPORSCSLSTPEAGALHVLPPARAEPQRLSFS
 FGDRLADLCVOAKASGILPYHSFLATLEDLSGFPSPHSIFSEVDASTOVLARI
 RYFPNMFLEKCHRGRLKRDASATLIDLPVLEHLPFOHRSDLVSRLPVLGSLKOG
 ECLSLAVLDLARMEQARQREGELKTYSKAKLPSIRDLTGSLGHDGILRVAGD
 RALPRVACQADRLSLMAKYIMDELRDPAGAEIVHGLVPTIRIDNOILEAEPG
 GGIAMTGGEOEVLQPCDFEIVDISIKQAPRGVAPRILEVARGCHPTLIDALINKLTG
 LPEALSPVALVDGFEPLTDSQHFCEKVPAPRILEVARGCHPTLIDALINKLTG
 GSRGSVYIARRSPQDPSPLTVCVONPLGPKGLIRSPGTLVLVGLSPHSLSL
 RELATCDGSLHYDVAVVLTSCIPRPEKSNLIVVGRSPPTSLVQPSQYL
 SÖMTFHKIPADSLNHERHNGSTKTYKRCGRHYVDGFAKTEVLKVDAAHKNM
 ESFLAASLMSQVSRYHLVLLHGVMAQCDSTWQEPVHIGALIMYLRKSHVAVLSIE
 LQVVOALAVNLVLELGLPHGNVSAKVLILARGAGSGVPTMPSALDPKAKIQTIED
 MLTDRIPWAVECELEAQTSLLEADKNGFQVWVEVSGVATMPSALDPKAKIQTIED
 RQOLPAPKTEALILIOCMAYEPVQPSFAVIRDNLSISDYELLSPDTGALAVQ
 RDGLNGQALVACODPTFEERHKLISQKNGFSGVSLCRDYPDLNHTGALVAVQ
 LQHSQEDQDRPQREIOTLTKALGSDFTYKVGVSYGGRPLRLVMEYLRISGLRDL
 LQHRARLIDASRLYSSQICGMETLSKRCVHNDLAARNLIVEAHVLIADGLAK
 LIPLDKDYVVRREPQSPITFWAPESLSDNIFSQSVMSGCVLYLFTYCDKSCSP
 SAELRMGCEEDYALCRLLELEEGRLPAPACPREVHELMKLCMAFSPODRPFS
 SALGPOLMWSGSRGCTHAFTAHPEKHLISLSFS"

BASE COUNT 746 a 1292 c 1168 g 858 t
 ORIGIN

Alignment Scores:
 Pred. No.: 8.02e-50 Length: 4064
 Score: 62.00 Matches: 96
 Percent Similarity: 97.96% Conservative: 0

Best Local Similarity: 97.96% Mismatches: 1
Query Match: 5.64% Indels: 2
DB: 9 Gaps: 0
US-09-397-967-16 (1-1099) x HSU09607 (1-4064)
OY 938 ArgAArgCysValHisArgAspLeuAlaAlaArgHisIleLeuValGluSerGluAlaHis 957
DB 2922 CGCCGCTGCGGACCGCGACGCTGGCCGGAACACTCTCTGCGAGACGCGGACAC 2981
OY 958 ValIleAlaAspPheGlyLeuAlaLysLeuLeuPro-LeuGlyLysAspTyrTyrVa 977
DB 2982 GTCAAGATTCGCTGACTGCGCTAGCTAACCTGCTGCGGCTTAC-AAAGACTACTACGT 3040
OY 977 ValArgGluProGlyGlnSerProIlePheTyrTyrAlaProGluSerLeuSerAspAs 997
DB 3041 GGTCCGCGACCGACGAGCCCATTTCTGCTGATGCCCGCAATCCCTCTCGGACAA 3100
OY 997 nIlePheSerArgGlnSerAspValTyrPheGlyValValLeuTyrGluLeuPheHem 1017
DB 3101 CATCTTCTCTGCGACGACGCTGAGACCTTGGGGGCTGCTGCTGACGAGCTTTCAC 3160
OY 1017 rTyrCysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetGly 1034
DB 3161 CTACTCGGACAAAGCTGACGCCCTCGCGCGAGTTCTGCGGATGATGGGA 3212
RESULT 10
MMU71201
LOCUS MMU71201 10326 bp DNA linear ROD 15-JAN-1997
DEFINITION Mus musculus protein tyrosine kinase JAK3 gene, complete cds.
ACCESSION U71201
VERSION 0.1
KEYWORDS GI:1778748
SOURCE Mus musculus.
ORGANISM Mus musculus.
REFERENCE 1 (bases 1 to 10326)
AUTHORS Kumar, A., Toscani, A., Rane, S. and Reddy, E.P.
TITLE Structural organization and chromosomal mapping of JAK3 locus
JOURNAL Oncogene 13 (9), 2009-2014 (1996)
MEDLINE 97088640
PUBMED 8934348
REFERENCE 2 (bases 1 to 10326)
AUTHORS Kumar, A. and Reddy, E.P.
TITLE Direct Submission
JOURNAL Submitted (18-SEP-1996) Fels Institute, Temple University, 3307 N.
Broad St, Rm 339 AHB, Philadelphia, PA 19140, USA
FEATURES
SOURCE
1. 10326
/organism="Mus musculus"
/strain="129/J"
/db_xref="taxon:10090"
/chromosome="19"
/map="19p13.1-p13.2"
/join(92..287,441..564,740..851,963..1108,1317..1599,
2170..2289,2372..2529,2617..2728,3230..3413,3781..3902,
3998..4129,4596..4680,5022..5149,5642..5774,5905..6056,
6177..6327,6856..6995,7074..7260,7427..7551,7629..7801,
8524..8641,8907..9017,9707..10212)
/product="JAK3"
/join(104..287,441..564,740..851,963..1108,1317..1599,
2170..2289,2372..2529,2617..2728,3230..3413,3781..3902,
3998..4129,4596..4680,5022..5149,5642..5774,5905..6056,
6177..6327,6856..6995,7074..7260,7427..7551,7629..7801,
8524..8641,8907..9017,9707..9772)
/function="protein tyrosine kinase"
/codon_start=1
/product_start=1
/product_end="JAK3"
/protein_id="A040917.1"
/db_xref="GI:1778748"
/translation="MAPSPEPPLIPORSCSSSSAGALHVLPPRGCGPPORLSFS
FGDYLAEDLCVRAKACGLIPYHSLPALATEDFSCPPSHIFCIEDVDIVLYRL

REYPPDFGLETCRFGRLKRDLTSAIDLHYLEHLFAOHRSDLYSGRLPYGLSMKEOG
EFLSLAVLDLAQMAREOORPGLKTVSKACIPSLRHRSDLYSGRLPYGLSMKEOG
LALRWSPARPTATAGGVYDILRLPAATTEFFRGGLGAGEEPGLVTRRRIRRV
MSGDOELFOTFCOPPELVYDSTIKOPTCGSGREHRLVTVMRGGHILEAFGLPEAL
SFVALVDSTFRLICDSRHRYFCKEVAAPRLLEEBELCHGPTITLDFALHKLAAALPG
TYILRSPODYDFLLACVOTPLGPGKGLITRQDGSAGFSLVSGSPPEACGTGCO
GNSGLRVDGALNLITSCAARPEKSNLIVRRGCTPAPAPGSPSCALNOLSFHT
IPTDLSHEINLGHSGSFPIKFRGRREVVDDEPDSVLLKVMDSRRHNCSESTLEAS
LMSOVYPHLVILHGVCMAGDSIMVDFEYVIGATDYLRLKRGHLYASMKLOVTKOLA
VALNLEDDKGLPHGNVARSARVILARBEQDDNPPFIKUSDVSPVLSLEMLDRIPL
VAPRCLOEACITCIEAKKRGRTATWVFSGGEPAHITSLSEPAKLFEYEDOGLPALK
WTELACITGCMATIDPGRRPSFRALILDLNGLITSDVELIDPTEPGLVAVKLOHSGPPO
OLVACQDPALFEERHLYKISLCKNGNSVELCRYPDLNTPGLVAVKLOHSGPPO
DFQREIOLIKALHSDFTIVRGVSGPGSGSLVMEYLPSCGLRDFLOHRAALHT
VRLLEFAMOLCKMGEYLGARCVHRLDAANILVSEFAHYKINDPGLAKLIPGKDY
VVRPGSPFIEMVAPELSLSPNFSROSDVMSFGVYLETGYCDKSCSRAEFLRMWG
PERGPEPLCRULEFLACGRRLPPPRCTPFVQLOLQVAPSTTTQSPAP"
BASE COUNT 2088 a 3007 c 2880 g 2300 t 51 others
ORIGIN
Alignment Scores:
Pred. No.: 1.69e-49 Length: 10326
Score: 62.00 Matches: 62
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.64% Indels: 0
DB: 10 Gaps: 0
US-09-397-967-16 (1-1099) x MMU71201 (1-10326)
OY 1 MetaIaProProSerGluGluThrProIleIleProGlnArgSerGlySerSerLeuSerSer 20
DB 104 ATGACACCTCCAGATGAGAGACACCTCTGATCCCTGAGCGCTCTGACACCTCTATCC 163
OY 21 SerIuAlaGlyAlaLeuHsValLeuLeuProProArgGlyProGlyProProGlnArg 40
DB 164 TCAGAGGACAGAGAGCCCTGATGCTCTCTCCGCCGAGGACCTGCCCTCCAGGCA 223
OY 41 LeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuProProArgGlyProGlyProGlnArg 60
DB 224 TTGTCTTCTCTTTGGGAGCTACTGCTGAGAGATTATGTGTGACACTGCCAAGGCC 283
OY 61 CysGly 62
DB 284 TGTGCT 289
RESULT 11
HSU70065
LOCUS HSU70065 13562 bp DNA linear PRI 27-JAN-1997
DEFINITION Human JAK3 gene, complete cds.
ACCESSION U70065
VERSION 0.1
KEYWORDS GI:1800224
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
REFERENCE 1 (bases 1 to 13562)
AUTHORS Bosse, A., O'Shea, J.T., Blake, T.B., Modi, W., Lal, B.K., Davis, J.,
Riedy, M.C., Dutra, A.S., and Johnston, J.A.
TITLE Genomic sequence, organization, and chromosomal localization of
human JAK3
JOURNAL Genomics 37 (1), 57-61 (1996)
MEDLINE 97079659
PUBMED 8921370
REFERENCE 2 (bases 1 to 13562)
AUTHORS Riedy, M.C., Dutra, A.S., Blake, T.B., Modi, W., Lal, B.K., Davis, J.,
Bosse, A., O'Shea, J.T., and Johnston, J.A.
TITLE Direct Submission
JOURNAL Submitted (09-SEP-1996) NIAHS, National Institutes of Health, 9000
Rockville Pike, Bethesda, MD 20892, USA
FEATURES
SOURCE
1. 13562
/organism="Homo sapiens"
/strain="T" (T) (1996)
/db_xref="taxon:9606"
/chromosome="9"
/map="9p24.3"
/join(104..287,441..564,740..851,963..1108,1317..1599,
2170..2289,2372..2529,2617..2728,3230..3413,3781..3902,
3998..4129,4596..4680,5022..5149,5642..5774,5905..6056,
6177..6327,6856..6995,7074..7260,7427..7551,7629..7801,
8524..8641,8907..9017,9707..9772)
/function="protein tyrosine kinase"
/codon_start=1
/product_start=1
/product_end="JAK3"
/protein_id="A040917.1"
/db_xref="GI:1778748"
/translation="MAPSPEPPLIPORSCSSSSAGALHVLPPRGCGPPORLSFS
FGDYLAEDLCVRAKACGLIPYHSLPALATEDFSCPPSHIFCIEDVDIVLYRL

[illegible]

intron	number=7 /gene="JAK3"
exon	number=7 4264..4450 /gene="JAK3"
intron	number=8 4451..6302 /gene="JAK3"
exon	number=8 6303..6430 /gene="JAK3"
intron	number=9 6431..6629 /gene="JAK3"
exon	number=9 6630..6761 /gene="JAK3"
intron	number=10 6762..7417 /gene="JAK3"
exon	number=10 7418..7503 /gene="JAK3"
intron	number=11 7504..7917 /gene="JAK3"
exon	number=11 7918..8044 /gene="JAK3"
intron	number=12 8045..8807 /gene="JAK3"
exon	number=12 8808..8939 /gene="JAK3"
intron	number=13 8940..9018 /gene="JAK3"
exon	number=13 9019..9171 /gene="JAK3"
intron	number=14 9172..9301 /gene="JAK3"
exon	number=14 9302..9452 /gene="JAK3"
intron	number=15 9453..10839 /gene="JAK3"
exon	number=15 10839..10978 /gene="JAK3"
intron	number=16 10979..11058 /gene="JAK3"
exon	number=16 11059..11256 /gene="JAK3"
intron	number=17 11257..12008 /gene="JAK3"
exon	number=17 12009..12125 /gene="JAK3"
intron	number=18 12126..12399 /gene="JAK3"
exon	number=18 12400..13562 /gene="JAK3"
	number=19

AC073750 223734 bp DNA linear HTG 29-JUN-2000
LOCUS 34800
DEFINITION Mus musculus clone RP23-330D8, WORKING DRAFT SEQUENCE, 28 unordered
pieces.
AC073750
AC073750.1 GI:8810367
HTG: HTGS_PHASE1; HTGS_DRAFT.
VERSION
KEYWORDS
SOURCE
ORGANISM
Mus musculus.
Mammalia; Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Eukaryota; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1. (bases 1 to 223734)
DOE Joint Genome Institute.
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Submitted (29-JUN-2000) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: <http://www.jgi.doe.gov>

Project Information
Center Project Name: 1863672
Center clone name: RP23-330D8

Summary Statistics
Consensus quality: 196566 bases at least Q40
Consensus quality: 208241 bases at least Q30
Consensus quality: 210965 bases at least Q20
Estimated insert size: 208000: agarose-fp estimation
Estimated insert size: 221034: sum-of-coverage estimation
Quality coverage: 8.01 in Q20 bases; agarose-fp estimation
Quality coverage: 7.54 in Q20 bases; sum-of-coverage estimation.
NOTE: This is a 'working draft' sequence. It currently
* consists of 28 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 1027: contig of 1027 bp in length
1127: gap of unknown length
1128 1127: contig of 1034 bp in length
2162 2261: gap of unknown length
2262 3709: contig of 1448 bp in length
3710 3809: gap of unknown length
3810 4864: contig of 1055 bp in length
4865 4964: gap of unknown length
4965 5985: contig of 1021 bp in length
5986 6085: gap of unknown length
6086 7691: contig of 1606 bp in length
7692 7791: gap of unknown length
7792 10317: contig of 2526 bp in length
10318 10417: gap of unknown length
10418 11756: contig of 1339 bp in length
11757 11856: gap of unknown length
11857 13323: contig of 1367 bp in length
13324 13324: gap of unknown length
13324 16114: contig of 2791 bp in length
16115 16214: gap of unknown length
16215 20778: contig of 4564 bp in length
20779 20878: gap of unknown length
20879 23553: contig of 2675 bp in length
23554 23653: gap of unknown length
23654 26667: contig of 3014 bp in length
26668 30222: gap of unknown length
30223 30322: gap of unknown length

30323 34899: contig of 4577 bp in length
34900 34989: gap of unknown length
35000 39609: contig of 4610 bp in length
39610 39709: gap of unknown length
39710 46418: contig of 6709 bp in length
46419 46518: gap of unknown length
46519 50029: contig of 3511 bp in length
50030 50129: gap of unknown length
50130 56982: contig of 6853 bp in length
56983 57082: gap of unknown length
57083 66435: contig of 9353 bp in length
66436 66535: gap of unknown length
66536 73174: contig of 6639 bp in length
73175 73274: gap of unknown length
73275 87407: contig of 14133 bp in length
87408 87507: gap of unknown length
87508 100238: contig of 12731 bp in length
100239 100338: gap of unknown length
100339 119196: contig of 18858 bp in length
119197 119296: gap of unknown length
119297 140119: contig of 20823 bp in length
140120 140219: gap of unknown length
140220 169227: contig of 29008 bp in length
169228 169327: gap of unknown length
169328 198979: contig of 29652 bp in length
198980 199079: gap of unknown length
199080 223734: contig of 24655 bp in length.
Location/Qualifiers
1. 223734
source
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="RP23-330D8"
/clone_11b="RP23 mouse BAC library 23"
BASE COUNT 57507 a 53185 c 54134 g 56203 t 2705 others
ORIGIN

Alignment Scores:
Pred. NO.: 1.98e-48 Length: 223734
Score: 62.00 Matches: 62
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.64% Indels: 0
DB: 2 Gaps: 0

US-09-397-967-16 (1-1099) x AC073750 (1-223734)

QY 1 MetaIapProserGluGluThrProLeuIleProGlnArgSerCysSerLeuSerSer 20
|||||
Db 124160 ATGGACCTCCAGAGGAGACACACTGTGATCCCTGACGCTTGCAGCCTCATCC 124219

QY 21 SerGluAlaGlyAlaLeuHisValLeuLeuProProArgGlyProGlyProProGlnArg 40
|||||
Db 124220 TCAGAGCGCAGGAGCCCTGCATGCTCTCTCCCTCCCGGAGCCTGGGCTCCCGCAGCGA 124279

QY 41 LeuSerPheSerPheGlyAspTyrLeuAlaGlnAspLeuCysValArgAlaAlaAlaAla 60
|||||
Db 124280 TTGTATTCTCTTTGGGACTACTTGCTGCTGAGATTATGTGTGCGACGCTGCCAAGGCC 124339

QY 61 CysGly 62
|||||
Db 124340 TGTGCT 124345

RESULT 14
LOCUS AX203086 795 bp DNA linear PAT 30-AUG-2001
DEFINITION Sequence 7 from Patent WO0152892.
ACCESSION AX203086
VERSION AX203086.1 GI:15392444
KEYWORDS
SOURCE

ORGANISM human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1 (bases 1 to 795)
 AUTHORS Vassios, G.
 TITLE Jak/stat pathway inhibitors and the uses thereof
 JOURNAL Patent: WO 0152892-A 7 26-JUL-2001;
 GENZYME CORPORATION (US)
 FEATURES Location/Qualifiers
 source 1. 795
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 BASE COUNT 156 a 255 c 226 g 158 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1.14e-39 Length: 795
 Score: 51.00 Matches: 51
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 4.64% Indels: 0
 DB: Gaps: 0
 US-09-397-967-16 (1-1099) x AX203086 (1-795)
 QY 838 PROASPGINGINARGASPHGGINARGIUGIIEGIIIEULYALALEUHSISERASP 877
 DB 455 CCAGACACAGACAGAGGACTTTCAGCGGAGATTGATCTCAAGCACTGCACAGTGAT 514
 QY 878 PHEIIIEVALLYSTRARGLYVALSERTRYGLYPROGLYARVAGLISERLEUARGLEUVAL 897
 DB 515 TTCATTGTCACATGATGCTGCTGCTACCTTATGCGCCGCGCCAGACGCTCGCGCTGCTC 574
 QY 898 MGLTIUTYLRLEUPROSERGLYCYSLEUARGASP 908
 DB 575 ATGAGACTACCTGCCAGCGGCTGCTGCCGCAC 607
 RESULT 15
 LOCUS AC103459/c
 DEFINITION Rattus norvegicus clone CH230-30H1, *** SEQUENCING IN PROGRESS ***
 68 unordered pieces.
 AC103459 159623 bp DNA linear HTG 12-JUL-2002
 AC103459
 AC103459.3 GI:21728723
 HTG: HTGS_PHASE1.
 SOURCE Norway rat.
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 159623)
 Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-osman, F.R., Allen, C.,
 Alsbrooks, S.L., Amaralunge, H.C., Are, J.R., Ayale, M., Banks, T.,
 Barbarta, J., Benton, J., Blimage, K., Blankenburg, K., Bonnin, D.,
 Bouck, J., Bowie, S., Brileva, M., Brown, E., Brown, M., Bryant, N.P.,
 Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,
 Carron, T.F., Carter, M., Cavazos, S.R., Checko, J., Chavez, D.,
 Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,
 Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,
 Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,
 Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,
 Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,
 Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,
 Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,
 Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,
 Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,
 Harris, C., Harris, K., Hart, M., Haylak, P., Hawes, A., Hernandez, J.,
 Hernandez, O., Hodgson, A., Hognes, M., Holloway, C., Hollins, B.,
 Homel, F., Howard, S., Huber, J., Hulik, S., Hume, J., Jackson, L.E.,
 Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,
 Karlsson, E., Kelly, S., Khan, U., King, L., Korvan, J., Kovar, C.,
 Kretovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,
 Li, D., Li, Z., Lichtarge, O., Liu, C., Liu, J., Liu, W., Louseged, H.,
 Lozdo, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,
 Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,
 Massey, E., Mawhney, E., McLeod, M.P., Meador, M., Mel, G., Metzker, M.,

REFERENCE 2 (bases 1 to 159623)
 AUTHORS Minner, G., Minner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,
 Moser, N., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,
 Nguyen, N., Nickerson, E., Nwokenkwo, S., Ogun, M., Okunodu, G.,
 Otagunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,
 Peters, L., Pickens, R., Primus, E., Pu, L.L., Qulles, M., Ren, X.,
 Rivas, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savery, G.,
 Scherer, S., Scott, G., Shen, H., Shoohtari, N., Sisson, I.,
 Sodergren, E., Sonalke, T., Sparks, A., Stanley, H., Stone, H.,
 Sutton, A., Swalek, A., Tabor, P., Tameris, A., Tameris, K., Tang, H.,
 Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
 Usmani, K., Vasquez, J., Vera, Y., Villalón, D., Vinson, R., Wang, Q.,
 Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
 Williams, G., Williamson, A., Wleczek, R., Wooden, S., Worley, K.,
 Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
 Weinstein, G., and Glbbs, R.
 Direct Submission
 Unpublished
 2 (bases 1 to 159623)
 Worley, K.C.
 Direct Submission
 Submitted (25-NOV-2001) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Jul 11, 2002 this sequence version replaced gi:17973301.
 COMMENT
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: http://www.hgsc.bcm.tmc.edu/
 Contact: hgsc-help@bcm.tmc.edu
 Project Information
 Center project name: CH230-30H1
 Center clone name: CH230-30H1
 Sequencing vector: Plasmid
 Chemistry: Dye-terminator Big Dye: 100% of reads
 Assembly program: Phrap: version 0.990329
 Consensus quality: 104279 bases at least Q40
 Consensus quality: 108242 bases at least Q30
 Consensus quality: 111035 bases at least Q20
 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank/draft_data.html).
 * NOTE: This is a "working draft" sequence. It currently
 * consists of 68 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 1 1035: contig of 1035 bp in length
 1036 1135: gap of unknown length
 1136 2180: contig of 1045 bp in length
 2181 2280: gap of unknown length
 2281 3531: contig of 1251 bp in length
 3532 3631: gap of unknown length
 3632 4743: contig of 1112 bp in length
 4744 4843: gap of unknown length
 4844 6313: contig of 1470 bp in length
 6314 7488: gap of unknown length
 7489 7588: contig of 1075 bp in length
 7589 7896: gap of unknown length
 7896 8895: contig of 1307 bp in length
 8896 8995: gap of unknown length
 8996 10362: contig of 1367 bp in length
 10363 11482: gap of unknown length
 11482: contig of 1020 bp in length

```

* 11483 11582: gap of unknown length
* 11583 12646: contig of 1064 bp in length
* 12647 12746: gap of unknown length
* 12747 14137: contig of 1391 bp in length
* 14138 14237: gap of unknown length
* 14238 15586: contig of 1349 bp in length
* 15587 15686: gap of unknown length
* 15687 16837: contig of 1151 bp in length
* 16838 16937: gap of unknown length
* 16938 18089: contig of 1152 bp in length
* 18090 18189: gap of unknown length
* 18190 19727: contig of 1538 bp in length
* 19728 19827: gap of unknown length
* 19828 21028: contig of 1201 bp in length
* 21029 22659: contig of 1541 bp in length
* 22660 22770: gap of unknown length
* 22770 24473: contig of 1704 bp in length
* 24474 24573: gap of unknown length
* 24574 25744: contig of 1171 bp in length
* 25745 25844: gap of unknown length
* 25845 27178: contig of 1334 bp in length
* 27179 27278: gap of unknown length
* 27279 28590: contig of 1312 bp in length
* 28591 28690: gap of unknown length
* 28691 31032: contig of 2342 bp in length
* 31033 31132: gap of unknown length
* 31133 32877: contig of 1745 bp in length
* 32878 34548: gap of unknown length
* 32978 34648: contig of 1571 bp in length
* 34549 36112: gap of unknown length
* 36113 36212: gap of unknown length
* 36213 37497: contig of 1285 bp in length
* 37498 37597: gap of unknown length
* 37598 39794: contig of 2197 bp in length
* 39795 39894: gap of unknown length
* 39895 41297: contig of 1403 bp in length
* 41298 41397: gap of unknown length
* 41398 42848: contig of 1451 bp in length
* 42849 42948: gap of unknown length
* 42949 44514: contig of 1566 bp in length
* 44515 44614: gap of unknown length
* 44615 46380: contig of 1766 bp in length
* 46381 46480: gap of unknown length
* 46481 47728: contig of 1248 bp in length
* 47729 47828: gap of unknown length
* 47829 49688: contig of 1860 bp in length
* 49689 49788: gap of unknown length
* 49789 51617: contig of 1829 bp in length
* 51618 51717: gap of unknown length
* 51718 53106: contig of 1389 bp in length
* 53107 53206: gap of unknown length
* 53207 54687: contig of 1481 bp in length
* 54688 54787: gap of unknown length
* 54788 56597: contig of 1810 bp in length
* 56598 56697: gap of unknown length
* 56698 58818: contig of 2121 bp in length
* 58819 60329: gap of unknown length
* 60330 60429: contig of 1411 bp in length
* 60430 61637: gap of unknown length
* 61638 61737: contig of 1208 bp in length
* 61738 63760: gap of unknown length
* 63761 63860: contig of 2023 bp in length
* 63861 65870: gap of unknown length
* 65871 65970: contig of 2010 bp in length
* 65971 67818: gap of unknown length
* 67819 67918: contig of 1848 bp in length
* 67919 70369: gap of unknown length
* 70370 70469: contig of 2451 bp in length
* 70470 73659: gap of unknown length
* 73660 73759: gap of unknown length

```

```

* 73760 76872: contig of 3113 bp in length
* 76873 76873 76872: gap of unknown length
* 76873 76973 76872: gap of unknown length
* 76973 79979: contig of 3007 bp in length
* 79980 80079: gap of unknown length
* 80080 83397: contig of 3318 bp in length
* 83398 83497: gap of unknown length
* 83498 86200: contig of 2703 bp in length
* 86201 86301 86200: gap of unknown length
* 86301 89469: contig of 3169 bp in length
* 89470 89569: gap of unknown length
* 89570 91678: contig of 2109 bp in length
* 91679 91778: gap of unknown length
* 91779 95960: contig of 4182 bp in length
* 95961 96061 95960: gap of unknown length
* 96061 98743 96061: gap of unknown length
* 98743 98843 98742: contig of 2682 bp in length
* 98843 101456: gap of unknown length

```

Alignment Scores:

```

Pred. No.: 7.5e-37 Length: 159623
Score: 50.00 Matches: 50
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.55% Indels: 0
DB: 2 Gaps: 0

```

US-09-397-967-16 (1-1099) x AC103459 (1-159623)

```

QY 13 GlnAGSerCysSerLeuSerSerSerGluAlaGlyAlaLeuHisValLeuProPro 32
DB 134614 CAGCCCTCTTGAGCCTCTCAACCTCAGACAGAGCCCGCATGCTCTCTCC 134555

```

```

QY 33 ArgGlyProGlyProGlyAlaArgLeuSerPheSerPheGlyAspTyrLeuAlaLysP 52
DB 134554 CCGGACCTGGGCTCCCGACGAGCTGATCTCTTTGGGACTACTGGCTAGGAC 134495

```

```

QY 53 LeuCysValArgAlaAlaLysAlaCysGly 62
DB 134494 CTGTGTGTGGGGCTGCCAGAGCCGTGTGTGT 134465

```

```

RESULT 16
LOCUS AX203080 448 bp DNA linear PAT 30-AUG-2001
DEFINITION Sequence 1 from Patent WO0152892.
ACCESSION AX203080
VERSION AX203080.1 GI:15392437
KEYWORDS
SOURCE human
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 448)
AUTHORS Vassios G.
TITLE Jnk/stat pathway inhibitors and the uses thereof
JOURNAL Patent: WO 0152892-A 1 26-JUL-2001;
GENZYME CORPORATION (US)
FEATURES
source 1..448
Location/Qualifiers

```

```

BASE COUNT 79 a 150 c 127 g 92 t
ORIGIN

```

Alignment Scores:

```

Pred. No.: 3.81e-29 Length: 448
Score: 40.00 Matches: 74
Percent Similarity: 97.37% Conservative: 0
Best Local Similarity: 97.37% Mismatches: 1
Query Match: 3.64% Indels: 2
DB: 6 Gaps: 0

```

US-09-397-967-16 (1-1099) x AX203080 (1-448)

QY 938 ArgAGCysValHtsArgAspLeuAlaAlaArgAsnIleLeuValGlnSerGluAlaHis 957
 Db 222 GCGCGCTGCTGCACCGCGACCTGGCCGCCGAAACATCTCTGGAGAGCGAGGCACAC 281
 QY 958 ValIysIleAlaAspPheGlyLeuAlaLysLeuLeuPro-IeuGlyLysAspTyrTyrVa 977
 Db 282 GTCAGATGCTGCTGACTTCGGCTAGCTGCTGCCGCTTAC-AAAGACTACTACGT 340
 QY 977 VALATGAGLuproGlyGlnSerProIlePheTyrTyrAlaProGlnSerLeuSerAspAs 997
 Db 341 GGTCCCGAGGCCAGCGCCAGCCCATTTCTGTGTATGCCCGCGAATCCCTTCGACAA 400
 QY 997 nilePheSerArgGlnSerAspAlaTrrpSerPheGlyValLeu 1012
 Db 401 CATCTTCTCTCCGCACTGACAGCTGTGAGACTTCGGGGTCTGCTCTG 446
 RESULT 17
 AF513860
 LOCUS DEFINITION Homo sapiens Janus kinase 3 (a protein tyrosine kinase, leukocyte)
 (JAK3) gene, complete cds.
 ACCESSION AF513860
 VERSION AF513860.1 GI:21263105
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE Rieder,M.J., Armet,T.Z., Carrington,D.P., Ozuna,M., Kuldanek,S.A.,
 AUTHORS Rajkumar,N., Totn,E.J., Yl.Q. and Nickerson,D.A.
 TITLE Direct Submission
 JOURNAL Submitted (20-MAY-2002) Genome Sciences, University of Washington,
 1705 NE Pacific, Seattle, WA 98195, USA
 COMMENT To cite this work please use: SeattleSNPS, NHBLI HL66682 Program
 for Genomic Applications, UW-FHCRC, Seattle, WA (URL:
 http://pga.gs.washington.edu).
 FEATURES
 source location/Qualifiers
 1..21323
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 repeat_region 1..214
 /rpt_family="Alu"
 variation 290
 /rpt_type-dispersed
 /frequency="0.13"
 /replace="a"
 variation 370
 /frequency="0.02"
 /replace="c"
 variation 575
 /frequency="0.33"
 /replace="a"
 variation 622
 /frequency="0.01"
 /replace="a"
 repeat_region 673..965
 /rpt_family="Alu"
 /rpt_type-dispersed
 variation 681
 /frequency="0.02"
 /replace="a"
 variation 784
 /frequency="0.13"
 /replace="a"
 variation 925
 /frequency="0.33"
 /replace="g"
 variation 970
 /frequency="0.01"
 /replace="c"
 variation 994
 /frequency="0.96"

repeat_region /replace=""
 1019..1270
 /rpt_family="Alu"
 /rpt_type-dispersed
 variation 1039
 /frequency="0.04"
 /replace="c"
 variation 1051
 /frequency="0.01"
 /replace="a"
 variation 1061
 /frequency="0.06"
 /replace="c"
 variation 1064
 /frequency="0.04"
 /replace="c"
 repeat_region 1273..1336
 /rpt_family="Alu"
 /rpt_type-dispersed
 repeat_region 1337..1630
 /rpt_family="Alu"
 /rpt_type-dispersed
 variation 1467
 /frequency="0.30"
 /replace="a"
 variation 1513
 /frequency="0.04"
 /replace="c"
 variation 1564
 /frequency="0.23"
 /replace="a"
 variation 1590..2091
 /note="Region not scanned for variation"
 1631..1706
 /rpt_family="Alu"
 /rpt_type-dispersed
 repeat_region 1966..2274
 /rpt_family="Alu"
 /rpt_type-dispersed
 gene <2359..20650
 /gene="JAK3"
 join(<2359..2557,2891..3014,3300..3411,3619..3764,
 4181..4475,5029..5151,5245..5402,6455..6566,7133..7319,
 8406..8533,8733..8864,9583..9667,10746..10873,
 11582..11714,11794..11945,12076..12226,13868..14007,
 14089..14278,14999..15123,15397..15569,16178..16295,
 16580..16690,19892..20650)
 /gene="JAK3"
 /product="Janus kinase 3 (a protein tyrosine kinase,
 leukocyte)"
 join(2374..2557,2891..3014,3300..3411,3619..3764,
 4181..4475,5029..5151,5245..5402,6455..6566,7133..7319,
 8406..8533,8733..8864,9583..9667,10746..10873,
 11582..11714,11794..11945,12076..12226,13868..14007,
 14089..14278,14999..15123,15397..15569,16178..16295,
 16580..16690,19892..20059)
 /gene="JAK3"
 /codon_start=1
 /product="Janus kinase 3 (a protein tyrosine kinase,
 leukocyte)"
 /protein_id="AA04860.1"
 /db_xref="GI:21263106"
 /translation="MAPSEETPLIPQSCSLSTFAGALHVLPAKGPGRPORLSFS
 FGDHLAEDLCVOAKASGILPYHSFLATDLSQPPSHIFSVEASQVLLYRI
 RFYFPMWPGLEKCHRGRLKDLASATIDLVLEHLPQHRSLDVSGRPLVGLSLKDG
 ECLSLAVDLARAKRQAPRGELTQVSKCLPSSLDLIQGLSFVIRRIKTRVR
 RADRVAAQQAQDHSLSMAKYIMDLERLDAGAAEFHVLPGALGHDGLGLRAGD
 GGIAMVQGEQEVLIQPCDPEIIVDISIKQAPVGAEGEHLVTRTDQIIEAEFG
 LPEALSFVALVDGYERLTIDSOHFCKEVAAPRLLEEVQEOCHGPTLDPAIINKKPG
 GSRPGSVLVRSPDQDFSLTVQVONPLGAPPKGLIIRSPGTGPIIVGLSRPHSL
 RELLATQMDGGILHVGSAVATLTSCTIPRKESNLIVDGRGSPPTSLVQOSQYL
 SMTFHKTRADSLTEHNEHNGHSFTKIRGCHREHVEVDGARKTEVLKAKMDKHKNCM
 ESTLEASLMSQVSTRHLVLHGVCAAGDSTWQEVHIGALIDMYLIRKGIHLVPSWK


```
/product="INL3_HUMAN [AA 1- 65]"
/protein_id="AAD22740.1"
/db_xref="GI:4558770"
/translation="MDRLPAMALVLLGPAIVFALGPAPTPENREKICGHHFVALVR
VCGGPRMSTERRPATGCDRE"
repeat_region complement(1486, .1786)
repeat_region /rpt_family="AluSx"
repeat_region complement(1803, .1970)
repeat_region /rpt_family="AluSg/x"
repeat_region complement(1975, .2275)
repeat_region /rpt_family="Aluub"
repeat_region complement(2279, .2444)
repeat_region /rpt_family="Aluuo"
repeat_region complement(2447, .2747)
repeat_region /rpt_family="AluSx"
repeat_region complement(2748, .2877)
repeat_region /rpt_family="Aluuo"
repeat_region 3052, .3357
repeat_region /rpt_family="AluSx"
repeat_region complement(3365, .3412)
repeat_region /rpt_family="MIR"
repeat_region complement(3773, .3846)
repeat_region /rpt_family="(GA)n"
repeat_region complement(3851, .4142)
repeat_region /rpt_family="AluSx"
repeat_region complement(4914, .4995)
repeat_region /rpt_family="MIR"
repeat_region 5386, .5686
repeat_region /rpt_family="AluY"
repeat_region 5739, .5848
repeat_region /rpt_family="FLAM_A"
repeat_region 5870, .6170
repeat_region /rpt_family="AluSx"
repeat_region complement(5918, .24239)
gene /gene="JAK3"
/note="Janus kinase 3 (a protein tyrosine kinase,
leukocyte)"
mRNA complement(join(5918, .6676,9874, .9984,10269, .10386,
10994, .11166,11440, .11564,12285, .12474,12556, .12695,
14337, .14487,14618, .14769,14849, .14981,15690, .15817,
16895, .16979,17698, .17829,18029, .18156,19243, .19429,
19996, .20107,21155, .21312,21406, .21528,22082, .22376,
22793, .22938,23146, .23257,23543, .23666,24000, .24239))
/gene="JAK3"
/product="Human JAK family protein tyrosine kinase (JAK3)
mRNA, complete cds."
repeat_region complement(6416, .6472)
repeat_region /rpt_family="MIR"
repeat_region complement(join(6509, .6676,9874, .9984,10269, .10386,
10994, .11166,11440, .11564,12285, .12474,12556, .12695,
14337, .14487,14618, .14769,14849, .14981,15690, .15817,
16895, .16979,17698, .17829,18029, .18156,19243, .19429,
19996, .20107,21155, .21312,21406, .21528,22082, .22376,
22793, .22938,23146, .23257,23543, .23666,24000, .24183))
/gene="JAK3"
/note="TYROSINE-PROTEIN KINASE JAK3; JANUS KINASE 3;
JAK-3; LEUKOCYTE JANUS KINASE; L-JAK"
/codon_start=1
/product="JAK3_HUMAN"
/db_xref="GI:4558771"
/protein_id="AAD22741.1"
/translation="MAPSEETPLIPQSCSLSTEGALHLLPARGPPORLSFS
FGDHLADLCVOAKASGITLVYHSLFALATEDLSCEPSSHIFSVEDASTOYLVI
RFFPNMFLGKCHRGFLRKDLASAIIDLPLEHLFPOHRSLOSGRLPGLSLKEGG
ECSLAVLILARNAREQARPGELIKTVSKACILPRLDLOLSLFTVRRIRRTVR
RALURVAAACQADRSLMAKIMDLERIDPAGAAFTFVIGLPGALGSDGJGLRVAAD
GGIATWQGEBOEVLOPCDFPEIVDISTKOAPRRGPAEHRILVYTRDNOILKEFPG
LPEALSFVALVDGEFLTLTDSOHECKEVAAPRLGLEVAQCHPITLDAIINKLKG
GSRGSYVLRRSPQDFDPSFLITVCVQNPGLPVGCLIRSPGTFLLVGLSRHSLS
RELATCMGDLGHVDAVATLTSCCIPRKEKSLIYVWGSHSPSTSLVOPOSOYL
SOMPFKIPADSLMEHNLGHSFTKLYRGCRHEVVDGEARKTEVILLKVDARKKNM
ESFLAESLMSQVSYRHLVLLHVCYMGAGDSTWQEPHLGLAIDMYLEKRRHLPAASK
LQYVKQLATVLTLEDKGLPHGNVSAKKVILLARAGADGSPFTKLSDPGVSPALVSL
```

```
repeat_region complement(6747, .7040)
repeat_region /rpt_family="AluSx"
repeat_region complement(7111, .7412)
repeat_region /rpt_family="AluSg"
repeat_region 7441, .7688
repeat_region /rpt_family="AluSg"
repeat_region complement(7985, .8020)
repeat_region /rpt_family="(CAAA)n"
repeat_region complement(8021, .8315)
repeat_region /rpt_family="Aluub"
repeat_region complement(8316, .8454)
repeat_region /rpt_family="Aluub"
repeat_region 8499, .8564
repeat_region /rpt_family="LINE2"
repeat_region 8643, .8902
repeat_region /rpt_family="Aluuo"
repeat_region complement(8929, .9224)
repeat_region /rpt_family="AluSx"
repeat_region 9243, .9536
repeat_region /rpt_family="Aluuo"
repeat_region complement(9560, .9753)
repeat_region /rpt_family="MIR"
repeat_region 11759, .12044
repeat_region /rpt_family="AluSx"
repeat_region complement(12074, .12154)
repeat_region /rpt_family="MIR"
repeat_region 13194, .13323
repeat_region /rpt_family="FLAM_A"
repeat_region 13337, .13638
repeat_region /rpt_family="AluSx"
repeat_region complement(13651, .13805)
repeat_region /rpt_family="Aluub"
repeat_region 13832, .14140
repeat_region /rpt_family="AluSx"
repeat_region complement(16047, .16337)
repeat_region /rpt_family="AluSx"
repeat_region complement(17180, .17480)
repeat_region /rpt_family="AluSx"
repeat_region complement(17487, .17609)
repeat_region /rpt_family="FLAM_C"
repeat_region complement(17910, .17976)
repeat_region /rpt_family="MIR"
repeat_region complement(18301, .18479)
repeat_region /rpt_family="AluSg/x"
repeat_region complement(18480, .18780)
repeat_region /rpt_family="AluSx"
repeat_region complement(19505, .19804)
repeat_region /rpt_family="AluSc"
repeat_region 20248, .20548
repeat_region /rpt_family="AluY"
repeat_region complement(20590, .20882)
repeat_region /rpt_family="Aluuo"
repeat_region 23362, .23462
repeat_region /rpt_family="MIR"
repeat_region complement(24063, .24094)
repeat_region /rpt_family="GC-rich"
repeat_region 24283, .24582
repeat_region /rpt_family="AluSx"
repeat_region 24583, .24608
repeat_region /rpt_family="(GAA)n"
repeat_region 24851, .24930
repeat_region /rpt_family="AluJ/FLAM"
repeat_region complement(24932, .25220)
repeat_region /rpt_family="AluSx"
repeat_region 25221, .25284
```

```
MLTDRIEMVAPECLREAOFTLSLEADKKWGFATVWEFSGTMTPIASLDPKAKLOFED
RQDLPAKWTLEALLIOOCMAVEPVORPSPRAVIRDLNLSISSPYELLSPTPCALAP
RDGLMNAQLYACADDPITFEERHKITISQLEKGNFVELCRTPLDDNGLAVAVQ
LOHSGPDQGRFOREIOTLKALHSDFTVKTGVSYGGRQSLRLVMEYLPSCGLRDL
QRHARARLDASRLLLYSQICKMEYLSRRCVHRDLARNILVSEAHVITADGLAK
LLPLDKDYVVRPEGQSPIFWYAPESLSDNIFSRQSPVWSGVVLYELFVYCDKCSG
SAEFLRMWGCEEDVYALCRLELLEEGORLPAAPACPAEVIHMLKLCMAVSPQORPSF
SALGPOLDMLWCSGRGCETHTAFTHPGKHHSLSFS"
```

```

/rpt_family="AluJo/ELAM"

Alignment Scores:
Pred. No.: 1.33e-26 Length: 41006
Score: 39.00 Matches: 39
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.55% Indels: 0
DB: 9 Gaps: 0

US-09-397-967-16 (1-1099) x AC007201 (1-41006)

QY 989 Tyra1aprog1uSerLeuSerAspAsn1IephSerArgInSerAspValTrpSerPhe 1008
Db 10385 TATGCCCGCATGCTCTCGACACATCTCTCTCGCCGCTCAGACGCTTGAGAGCTTC 10326

QY 1009 GlyValValLeuTyrGluLeuPheThrTyrCysAspLysSerCysSerProSerAla 1027
Db 10325 GGGGTCGTCTGTACGAGAGCTCTTCACCTACTGCGACAAAGCTCGACGCCCTCGGCC 10269

RESULT 19
LOCUS HSU08340 198 bp mRNA linear PRI 21-APR-1994
DEFINITION Human clone NTK16 tyrosine kinase mRNA, partial cds.
ACCESSION U08340
VERSION U08340.1 GI:473879
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Fuortes M.
TITLE Tumor Necrosis Factor and Adhesion: Joint Control of Neutrophil
JOURNAL Activation
Tnesis (1994) Cell Biology and Anatomy, Cornell University,
Graduate School of Medical Sciences
2 (bases 1 to 198)
AUTHORS Fuortes M.
TITLE Direct Submission
JOURNAL Submitted (05-APR-1994) Michele Fuortes, Cell Biology and
Anatomy/Medicine, Cornell University Medical College, 1300 York
Avenue, New York, NY 10021, USA
FEATURES
source
1. 198
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="NTK16"
/cell_type="neutrophil"
/tissue_type="blood"
<1..>198
/codon_start=1
/evidence_start=1
/evidence_end=experimental
/product="tyrosine kinase"
/protein_id="AA17743.1"
/db_xref="GI:473880"
/translation="DLAARNIVLESEAHVIAFDGLAKLLPLDKDYVVEPQSPLE
WYAPESIDNIFRSQDVMSEFG"
BASE COUNT 41 a 66 c 52 g 39 t
ORIGIN

Alignment Scores:
Pred. No.: 1.66e-26 Length: 198
Score: 37.00 Matches: 65
Percent Similarity: 97.01% Conservative: 0
Best Local Similarity: 97.01% Mismatches: 1
Query Match: 3.37% Indels: 2
DB: 9 Gaps: 0

US-09-397-967-16 (1-1099) x HSU08340 (1-198)

QY 944 AspleuAlaAaGAsn1IleLeuValGluSerGluAlaHisValLysIleAlaAspPhe 963
|||||

```

```

Db 1 GACCTGGCGCCCGCAACATCTCTGTGAGAGCGAGCCACACGTCAGATCGTCACTTC 60

QY 964 GlyLeuAlaLysLeuLeuPro-LeuGlyLysAspTyrTyrValValArgLupProGly 983
Db 61 GGCTTACTAAGCTGCGTCCGCTTGAC-AAAGACTACTACTGCTCGCGAGCCAGGCCA 119

QY 983 nserPro1IephThrPyra1aprog1uSerLeuSerAspAsn1IephSerArgInse 1003
Db 120 GAGCCCATTTTCTGTGATATGCGCCGATCCCTCTCGGACACATCTTCTCTGCCAGTC 179

QY 1003 rAspValTrpSerPhegly 1009
Db 180 AGACGTGTGAGACTTCGGA 198

RESULT 20
LOCUS E08798 165 bp RNA linear PAT 29-SEP-1997
DEFINITION cDNA encoding novel human tyrosine kinase.
ACCESSION E08798
VERSION E08798.1 GI:2176910
KEYWORDS JP 1995059569-A/1.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Sakano S.
TITLE NOVEL TYROSINE KINASE PEPTIDE AND DNA CODING THE SAME
JOURNAL ASAHI CHEM IND CO LTD
OS Homo sapiens (human)
PN JP 1995059569-A/1
PD 07-MAR-1995
PF 25-AUG-1993-GB-1993210404
PI SAKANO SEIJI
PC C12N15/09,C12N9/12;
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
FH key Location/Qualifiers
FT source
1. 165
/organism="Homo sapiens"
/cell_type="megakaryoblast"
/cell_line="U9-7"
mat_peptide
1. 165
/product="novel human tyrosine kinase"
Location/Qualifiers
source
1. 165
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 36 a 56 c 39 g 34 t
ORIGIN

Alignment Scores:
Pred. No.: 1.01e-20 Length: 165
Score: 31.00 Matches: 31
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.82% Indels: 0
DB: 6 Gaps: 0

US-09-397-967-16 (1-1099) x E08798 (1-165)

QY 973 LysAspTyrTyrValValArgLupProGlyInserPro1IephThrPyra1aprog1u 992
Db 73 AAAGACTACTAAGCTGCGTCCGAGCCAGCCAGCCCATTTTCTGTGATGCGCCGAA 132

QY 993 SerLeuSerAspAsn1IephSerArgInser 1003
Db 133 TCCTCTCGGACACATCTTCTCTGCCAGTCA 165
|||||

```

```

RESULT 21
LOCUS AF034576 3644 bp mRNA linear VRT 03-SEP-1998
DEFINITION Gallus gallus Janus tyrosine kinase (JAK) mRNA, complete cds.
ACCESSION AF034576
VERSION AF034576.1 GI:2645986
KEYWORDS
SOURCE
ORGANISM Gallus gallus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE
AUTHORS Sofer, L., Kampa, D. and Burnside, J.
TITLE Molecular cloning of a chicken JAK homolog from activated T cells.
JOURNAL Gene 215 (1), 29-36 (1998)
MEDLINE 98332716
PUBMED 9666067
REFERENCE
AUTHORS Sofer, L., Kampa, D. and Burnside, J.
TITLE Direct Submission
JOURNAL Submitted (14-NOV-1997) Animal Science, University of Delaware, 40
Townsend Hall, Newark, DE 19716, USA
FEATURES
source
1..3644
location/Qualifiers
/organism="Gallus gallus"
/sub_species="domesticus"
/db_xref="taxon:9031"
/cell_type="lymphocyte"
/tissue_type="spleen"
/notes="cells activated with concanavalin A"
1..3644
/gene="JAK"
91..3411
/gene="JAK"
/feature="JAK"
/feature="similar to JAK protein kinase"
/codon_start=1
/product="Janus tyrosine kinase"
/protein_id="AAC34195.1"
/db_xref="GI:2645987"
translation="MAPLGEETPLIGRSCSISSEPTGLQVLYTHRGHPAPPSAAT
LTFEGEYAEELCYHAACGVLPICHPALATIEDLSCYPPNHITVTVDASOY
VYRIFFEPNMGOGVHROPSPRPVADIPYIDLFKSSSDPFIAGMELASL
AGOECLSLAVLDMRIKERMOSKPEFVSVKTCIPEDLRQIQSHSLTKRIK
RVAOSLRMGSCRYDCCCLAKLILDLERLCCMAESFHAHPDADIALHVEPTDS
VSMSCVSESQHFCDPDIADVSIKQNSRGGVEMNIVLTITDNVLEPTTLR
EALSFMAVLDGYRLTTADAHYFCKEVAPRLLEDNMOCHGPTSEFAVKLKAAS
HPGLVLRSPDQEDSYLLTYCAETRSQDYKRLIRDEGGSFMLGIARFCSLOE
LLGTGCGGLAEGAHLLDTCCPLPREKSNLLIVRSCGCPRPNSPPAPRSPNOS
FKHIDPESLIGESLGGSPFTHIKYIKIRKDDDEPOTPVYLVMSHRCSESEL
EASISMSLSHKLHYLLHGVSLGKDSIMVQYTRHPLDLYLKNHSEGVYTSKIQ
VAKOLATALNLEDKITTHGNVSAKKVLTREGDAASSPFILNDPGVSTVLAK
WLVERIPVNAPECLSDPQSLAPADKMGFGATLWEIFSGMMPVSLPEPKLQFYS
RIOLPAPRMSLALIAQCIAPSRPFRFIIINDISLSDSELSPEPVVTR
ESGQYEHVAGHAPPEERHLYKISLKGNGSVELICQYDPLGDSGTGLAVKTL
QODSAKLEDFEREIOLSHLDEIVIKRGVCYSRGRRLRLMETLPODCLDYIQ
KNORHLEHRTLLYAMOCCKMEYLGARQCHRDRLASNLIVSETHVKIGDFLANL
LPQDKDYVVOPEQSPFVWYAPASLADNVSFASDVSFVLLYELFTYSNKRSPS
EERLHMGPEKPAQIICHLLELLLDLSRRLPVPCGPMIEVYAMLSQWAFASAPRTPT
ELARVLEALRGRTANG"
BASE COUNT 697 a 1173 c 1150 g 624 t
ORIGIN

```

Alignment Scores:

```

Pred. No.: 1.14e-18 Length: 3644
Score: 30.00 Matches: 30
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.73% Indels: 0
DB: 5 Gaps: 0

```

US-09-397-967-16 (1-1099) x AF034576 (1-3644)

```

Qy 813 PhcGlUGlUArghHisLeuLysTyrIleSerLeuLeuGlyLysGlyAsnPhcGlYserVal 832
Db 2536 TTCGAGAGAGAGCACCCTCAAGTACATCTGCTGGGAAAGGCAACTTGGAGCGTG 2595
Qy 833 GlULeuCySArgTyrAspProLeuGlyAsp 842
Db 2596 GACCTGTGCCGCTACGACCCGCTGGGTGAC 2625
RESULT 22
LOCUS AX203088 811 bp DNA linear PAT 30-AUG-2001
DEFINITION Sequence 9 from Patent WO0152892.
ACCESSION AX203088
VERSION AX203088.1 GI:15392446
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Vassios, G.
TITLE Jak/stat pathway inhibitors and the uses thereof
JOURNAL Patent: WO 0152892-A 9 26-JUL-2001;
GENZYME CORPORATION (US)
FEATURES
source
1..811
location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 149 a 254 c 226 g 182 t
ORIGIN

```

Alignment Scores:

```

Pred. No.: 3.22e-18 Length: 811
Score: 29.00 Matches: 29
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.64% Indels: 0
DB: 6 Gaps: 0

```

US-09-397-967-16 (1-1099) x AX203088 (1-811)

```

Qy 1006 TRPSPERPhcGlyValValLeuTyrGluLeuPhcThrTyrCysAspLysSerCysSerPro 1025
Db 168 TGGACCTTCGGGGGCGTCTCTACGAGCTCTTCACTGCGACAAAGCTGACGCC 227
Qy 1026 SerIaGluPhcLeuArcMetMetGly 1034
Db 228 TCGGCGCGAGTTCCTGCGGATGATGGA 254

```

```

RESULT 23
LOCUS AC103459 159623 bp DNA linear HTG 12-JUL-2002
DEFINITION Rattus norvegicus clone CH230-30H1, *** SEQUENCING IN PROGRESS ***
ACCESSION AC103459
VERSION AC103459.3 GI:21728723
KEYWORDS HTG; HTGS; PHASEL.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 159623)

```

REFERENCE
 Muzny, D.M., Adams, C., Adio-Oduola, B., All-ouman, F.R., Allen, C.,
 Alsbrooks, S.L., Amaralunga, H.C., Are, J.R., Ayale, M., Banks, T.,
 Barbara, J., Benton, J., Bimaga, K., Blankenburg, K., Bonnin, D.,
 Bouck, J., Bowls, S., Brileva, M., Brown, E., Brown, M., Bryant, N.P.,
 Buha, C., Burch, P., Burkett, C., Burrell, R.L., Byrd, N.C.,
 Carion, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,
 Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,
 Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,
 Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,
 Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,

Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,
 Eanharth, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,
 Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Franze, P.,
 Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,
 Gottlieb, O., H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,
 Harris, C., Harris, K., Hart, M., Havlik, P., Hawes, A., Hernandez, J.,
 Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, B.,
 Homsl, F., Howard, S., Huber, J., Huliy, S., Hume, J., Jackson, L.E.,
 Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,
 Karlsson, E., Kelly, S., Khan, L., King, L., Korvah, J., Kovar, C.,
 Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,
 Li, J., Li, Z., Licharge, O., Lieu, C., Liu, J., Liu, W., Loulseged, H.,
 Lozada, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,
 Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,
 Massey, E., Mawhney, E., McLeod, M.P., Meador, M., Mel, G., Metzger, M.,
 Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,
 Moser, M., Neal, D., Newton, J., Newton, S., Ogun, M., Okunolu, G.,
 Nguyen, N., Nickerson, E., Nwokenwo, S., Ogun, M., Okunolu, G.,
 Orangunye, N., Oyedero, R., Pace, A., Payton, B., Peery, J., Perez, L.,
 Peters, L., Pickens, R., Plimms, E., Pu, L.L., Quiles, M., Ren, Y.,
 Rivers, M., Rojas, A., Rojubokan, I., Rolfe, M., Ruiz, S., Saverly, G.,
 Scherer, S., Scott, G., Shen, H., Shooshitari, N., Sisson, I.,
 Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H.,
 Sutton, A., Syatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,
 Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
 Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wang, Q.,
 Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
 Williams, G., Williamson, A., Wleczek, R., Wooden, S., Worley, K.,
 Wu, C., Wu, Y., Wu, Y., Zhou, J., Zorrilla, S., Nelson, D.,
 Weinstein, G., and Gibbs, R.
 Unpublished
 Direct Submission
 2 (bases 1 to 159623)
 Worley, K.C.
 Direct Submission
 Submitted (25-NOV-2001) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 159623)
 Worley, K.C.
 Direct Submission
 Submitted (12-JUL-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Jul 11, 2002 this sequence version replaced gi:17973301.
 ----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GH00
 Center clone name: CH230-30H1
 ----- Summary Statistics
 Sequencing vector: Plasmid
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990329
 Consensus quality: 104279 bases at least Q40
 Consensus quality: 108242 bases at least Q30
 Consensus quality: 111035 bases at least Q20

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank/draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 68 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 * 1 1035: contig of 1035 bp in length
 * 1036 1135: gap of unknown length
 * 1136 2180: contig of 1045 bp in length
 * 2181 2280: gap of unknown length
 * 2281 3531: contig of 1251 bp in length
 * 3532 3632: gap of unknown length
 * 3632 4744: contig of 1112 bp in length
 * 4744 4844: gap of unknown length
 * 4844 6314: contig of 1470 bp in length
 * 6314 7488: gap of unknown length
 * 7488 7889: contig of 1075 bp in length
 * 7889 8895: gap of unknown length
 * 8895 8995: gap of unknown length
 * 8995 10363: gap of unknown length
 * 10363 10463: gap of unknown length
 * 10463 11582: contig of 1020 bp in length
 * 11582 12646: gap of unknown length
 * 12646 12647: contig of 1064 bp in length
 * 12647 14137: gap of unknown length
 * 14137 14237: contig of 1391 bp in length
 * 14237 14338: gap of unknown length
 * 14338 14586: contig of 1349 bp in length
 * 14586 15687: gap of unknown length
 * 15687 16937: contig of 1151 bp in length
 * 16937 16938: gap of unknown length
 * 16938 18089: contig of 1152 bp in length
 * 18089 18189: gap of unknown length
 * 18189 18190: contig of 1538 bp in length
 * 18190 19727: gap of unknown length
 * 19727 19828: contig of 1201 bp in length
 * 19828 21029: gap of unknown length
 * 21029 21129: gap of unknown length
 * 21129 22670: contig of 1541 bp in length
 * 22670 22770: gap of unknown length
 * 22770 24474: contig of 1704 bp in length
 * 24474 25744: gap of unknown length
 * 25744 25844: contig of 1171 bp in length
 * 25844 27178: gap of unknown length
 * 27178 27279: contig of 1334 bp in length
 * 27279 28590: gap of unknown length
 * 28590 28691: contig of 1312 bp in length
 * 28691 31032: gap of unknown length
 * 31032 31132: contig of 2342 bp in length
 * 31132 32878: gap of unknown length
 * 32878 32978: contig of 1745 bp in length
 * 32978 34548: gap of unknown length
 * 34548 34648: contig of 1571 bp in length
 * 34648 36112: gap of unknown length
 * 36112 36213: contig of 1464 bp in length
 * 36213 37497: gap of unknown length
 * 37497 37597: contig of 1285 bp in length
 * 37597 39794: gap of unknown length
 * 39794 39894: contig of 2197 bp in length
 * 39894 41297: gap of unknown length
 * 41297 41398: contig of 1403 bp in length
 * 41398 42848: gap of unknown length
 * 42848 42948: contig of 1451 bp in length
 * 42948 44514: gap of unknown length
 * 44514 44614: contig of 1566 bp in length
 * 44614 46380: gap of unknown length
 * 46380 46480: contig of 1766 bp in length
 * 46480 47728: gap of unknown length
 * 47728 47828: contig of 1248 bp in length
 * 47828 49688: gap of unknown length
 * 49688 49788: contig of 1860 bp in length
 * 49788 51617: gap of unknown length
 * 51617 51717: contig of 1829 bp in length
 * 51717 53206: gap of unknown length
 * 53206 53207: contig of 1389 bp in length
 * 53207 54687: gap of unknown length
 * 54687 54787: contig of 1481 bp in length
 * 54787 56597: gap of unknown length
 * 56597 56697: contig of 1810 bp in length
 * 56697 58818: gap of unknown length
 * 58818 58818: contig of 2121 bp in length

ACCESSION AX203082
VERSION AX203082.1 GI:15392439
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Euthera; Primates; Catarrhini; Homiidae; Homo.
TITLE Vasio, G.
JOURNAL Jak/stat pathway inhibitors and the uses thereof
PATENT: WO 0152892-A 3 26-JUL-2001;
GENZYME CORPORATION (US)
FEATURES
source Location/Qualifiers
1.109
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 22 a 36 c 34 g 15 t 2 others
ORIGIN

Alignment Scores:
Pred. No.: 3.4e-08 Length: 109
Score: 18.00 Matches: 18
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.64% Indels: 0
DB: 6 Gaps: 0

US-09-397-967-16 (1-1099) x AX203082 (1-109)

QY 945 leuAlaAaGaSnlleuValGluSerGluAlaHisVallylleAlaasp 962
|||||
56 TTGGCCGCCGAAACATCTCTGAGAGAGGACGACACGTCAGATGCTGAC 109

RESULT 26
AX203085 726 bp DNA linear PAT 30-AUG-2001
LOCUS
DEFINITION Sequence 6 from Patent W00152892.
ACCESSION AX203085
VERSION AX203085.1 GI:15392442
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Euthera; Primates; Catarrhini; Homiidae; Homo.
TITLE Vasio, G.
JOURNAL Jak/stat pathway inhibitors and the uses thereof
PATENT: WO 0152892-A 6 26-JUL-2001;
GENZYME CORPORATION (US)
FEATURES
source Location/Qualifiers
1.726
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 151 a 224 c 204 g 147 t
ORIGIN

Alignment Scores:
Pred. No.: 1.55e-07 Length: 726
Score: 18.00 Matches: 18
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.64% Indels: 0
DB: 6 Gaps: 0

US-09-397-967-16 (1-1099) x AX203085 (1-726)

QY 559 AsncysmetGluSerleuGluAlaAlaSerleuMetSerGlnValSerTyr 576
|||||
Db 635 AACGCAATGAGATCTCTGAGAGAGGAGCTGATGACCAAGTCTCTAC 688

RESULT 27
AF227557 495 bp mRNA linear PRI 01-JUN-2000
LOCUS

DEFINITION Macaca mulatta Janus kinase-3 (JAK3) mRNA, partial sequence.
AF227557
VERSION AF227557.1 GI:8132806
KEYWORDS
SOURCE Macaca mulatta.
ORGANISM Macaca mulatta
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Euthera; Primates; Catarrhini; Cercopitheidae;
TITLE Cercopitheidae; Macaca.
JOURNAL 1 (bases 1 to 495)
AUTHORS Attwood, J.
TITLE Cytokine Signal Transduction Genes from Rhesus Macaques
JOURNAL 2 (bases 1 to 495)
REFERENCE Arredondo, J.
AUTHORS Direct Submission
TITLE Submitted (24-JAN-2000) California Regional Primate Research
JOURNAL Center, University of California-Davis, One Shields Avenue, Davis,
CA 95616, USA
FEATURES
source Location/Qualifiers
1.495
/organism="Macaca mulatta"
/db_xref="taxon:9544"
/cell_type="peripheral blood mononuclear cells"
gene <1..>495
/gene="JAK3"
misc_feature <1..>495
/note="similar to Janus kinase-3; tyrosine kinase class I
receptor; contains frameshifts"
BASE COUNT 100 a 157 c 132 g 106 t
ORIGIN

Alignment Scores:
Pred. No.: 1.08e-06 Length: 495
Score: 17.00 Matches: 17
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.55% Indels: 0
DB: 9 Gaps: 0

US-09-397-967-16 (1-1099) x AF227557 (1-495)

QY 514 AsperleuGluTrpHisGluAsnleuGlyHisGlySerPheTrpIle 530
|||||
Db 376 GACAGCCTAGAGTGCGATGAGAGACCTGGCCATGATCTTCACCAAGATT 426

RESULT 28
HSU57096 466 bp mRNA linear PRI 18-JUN-1996
LOCUS
DEFINITION Human Janus kinase 3 (JAK3) mRNA, partial cds.
ACCESSION U57096
VERSION U57096.1 GI:1354824
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Euthera; Primates; Catarrhini; Homiidae; Homo.
TITLE Versky, J.W., Bach, E.A., Fang, Y.F., Yang, L., Randall, D.A. and
Fields, L.E.
JOURNAL Expression of Janus kinase 3 in human endothelial and other
non-lymphoid and non-myeloid cells
JOURNAL J. Biol. Chem. 271 (24), 13976-13980 (1996)
MEDLINE 96278845
PUBMED 8662778
REFERENCE 2 (bases 1 to 466)
AUTHORS Fields, L.E.
TITLE Direct Submission
JOURNAL Submitted (30-APR-1996) Larry E. Fields, Medicine, Washington
University School of Medicine, 660 S. Euclid Ave., Campus Box 8086,
St. Louis, MO 63110-1093, USA
FEATURES
source Location/Qualifiers


```

source
1. .466
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_type="vascular endothelial cells (HUVEC) obtained
from Clonetics Corporation, San Diego, CA"
gene
1. .466
/gene="jak3"
<1. .466
/gene="jak3"
/feature="phosphorylates Stat6, IRS1, IRS2, PI3K"
/feature="protein tyrosine kinase"
/codon_start=1
/product="janus kinase 3"
/protein_id="AAC50542.1"
/db_xref="GI:1354825"
/translation="GPGPRLSFSFGDHLAEDICVQAAKAGTIPVYHSPLALRDL
CWPPSHIFSYEDASTOVLIRIRIFPNMFGLEKRCRGLKDLASITLDPVLEHL
PAQHRSDLVGRLPVGLSLKEGELSLAVLDLARMARQAOAPPELLKAVSYK"
BASE COUNT      70 a      153 c      143 g      100 t
ORIGIN

Alignment Scores:
Pred. No.:      9.68e-06      Length:      466
Score:          16.00      Matches:      16
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    1.46%      Indels:      0
DB:             9      Gaps:      0

US-09-397-967-16 (1-1099) x HSU57096 (1-466)

OY 62 G1Y1LeuProValTYrHisSerLeuPheAlaLeuAlaThrGluAsp 77
Db 79 GGCATCTGCGCTGTACCACTCCTCTTGTCTGCGCAGCAGAGAC 126

RESULT 29
DRE5691      1967 bp      mRNA      linear      VRT 30-SEP-1998
LOCUS      DRE5691      1967 bp      mRNA      linear      VRT 30-SEP-1998
DEFINITION      Dario rerio mRNA for protein tyrosine kinase (jak2b), partial.
ACCESSION      AJ005691
VERSION      AJ005691.1      GI:3687399
KEYWORDS      jak2b gene; protein tyrosine kinase; signal transduction.
SOURCE      zebrafish.
ORGANISM      Dario rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Dario.
1 (bases 1 to 1967)
Oates A.C.
Direct Submission
Submitted (22-APR-1998) Oates A.C., Growth Regulation, Ludwig
Institute for Cancer Research, PO Royal Melbourne Hospital,
Parkville, VIC, 3050, AUSTRALIA
2 (bases 1 to 1967)
Oates A.C.
Thesis (1998) University of Melbourne, Australia
Location/Qualifiers
1. .1967
/organism="Dario rerio"
/db_xref="taxon:7955"
1. .1497
/gene="jak2b"
<1. .1497
/gene="jak2b"
/EC_number="2.7.1.12"
/function="signal transduction"
/codon_start=1
/product="protein tyrosine kinase"
/protein_id="CAA0675.1"
/db_xref="GI:3687400"
/db_xref="SPTREMBL:O93597"
/translation="SVNIIWKLEVAQKOLANAMLYLEKSLAHGVCANILLIREDR
ALGNTPIKISDPGISITVLPREILVERIFWVPEPCILDPKRNLSLADKMSFGTTLME

```

```

ICSGEOPLIAMNDSSKKHLPVENNHQLPARKWTELANLINSCHMEPEPTPRSKAIR
DLNLSFCDPDIYVESDIMPSSRAASTFNTGTFRNNEVQEEERHLIFLOQKGNG
SVEMCRYPDLDNTNGEYVAAYVKRKHSTTEHTRDEREIEILKSLQHENIYKGCYQ
AGRNRRLVMEYELPYGSLRDYLNKRNPRIDHOKLVHASQICMEYLAIRYIHRDL
ATRNILVESECRVKIGDFGLTKVLPQREYKKVKEPESPIFWAPSPSTESKFSVAS
DVSEGVLYELFTYSDKICSPTPVFLSMGDKGOGOTIYVHLIELKRGRLRPPQC
CPTMEFEMQCMNDPSLRNREKELARDLINDSSDADRYTQVPER"
BASE COUNT      563 a      429 c      498 g      476 t      1 others
ORIGIN

Alignment Scores:
Pred. No.:      3.06e-05      Length:      1967
Score:          16.00      Matches:      16
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    1.46%      Indels:      0
DB:             5      Gaps:      0

US-09-397-967-16 (1-1099) x DRE5691 (1-1967)

OY 1003 SerAspValTrpSerPheGlyValLeuTyrgluLeuPheThrTyr 1018
Db 1174 TCCGACGTGTGAGATTTCGAGTGTCTGTATGAACCTTTACATAC 1221

RESULT 30
ARI29823      3435 bp      DNA      linear      PAT 16-MAY-2001
LOCUS      ARI29823      3435 bp      DNA      linear      PAT 16-MAY-2001
DEFINITION      Sequence 21 from patent US 6187552.
ACCESSION      ARI29823
VERSION      ARI29823.1      GI:14117720
KEYWORDS
SOURCE      Unknown.
ORGANISM      Unknown.
REFERENCE      1 (bases 1 to 3435)
AUTHORS      Roberts,S.L. and Kaytes,P.S.
TITLE      Method for identifying inhibitors of JAK2/cytokine receptor binding
JOURNAL      Patent: US 6187552-A 21 13-FEB-2001;
FEATURES
source      Location/Qualifiers
1. .3435
/organism="unknown"
BASE COUNT      1157 a      598 c      733 g      947 t
ORIGIN

Alignment Scores:
Pred. No.:      4.78e-05      Length:      3435
Score:          16.00      Matches:      16
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    1.46%      Indels:      0
DB:             6      Gaps:      0

US-09-397-967-16 (1-1099) x ARI29823 (1-3435)

OY 1003 SerSpValTrpSerPheGlyValLeuTyrgluLeuPheThrTyr 1018
Db 3106 TCAGATGTTTGAGACTTGGAGTGTCTGTATGAACCTTTACATAC 3153

RESULT 31
AR048200      3495 bp      DNA      linear      PAT 29-SEP-1999
LOCUS      AR048200      3495 bp      DNA      linear      PAT 29-SEP-1999
DEFINITION      Sequence 2 from patent US 5821069.
ACCESSION      AR048200
VERSION      AR048200.1      GI:5970543
KEYWORDS
SOURCE      Unknown.
ORGANISM      Unknown.
REFERENCE      1 (bases 1 to 3495)
AUTHORS      Wilks,A.Frederick., Ziemlecki,A. and Harpur,A.
TITLE      Method for determining tyrosine kinase in a sample
JOURNAL      Patent: US 5821069-A 2 13-OCT-1998;
FEATURES      Location/Qualifiers

```

```

source 1..3495
/organism="unknown"
BASE COUNT 1121 a 688 c 794 g 892 t
ORIGIN

Alignment Scores:
Pred. No.: 4.85e-05 Length: 3495
Score: 16.00 Matches: 16
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.46% Indels: 0
DB: Gaps: 0

US-09-397-967-16 (1-1099) x AR048200 (1-3495)

OY 1003 SeraspValTTPSerPheGlyValValLeuTyrGluLeuPheThrTyr 1018
|||||
Db 2695 TCAGATGTGTGAGCTTGGAGTGTCTCTATACGAACTTTTCACATAC 2742

RESULT 32
AR068116 AR068116 3495 bp DNA linear PAT 29-SEP-1999
LOCUS Sequence 2 from patent US 5852184.
ACCESSION AR068116
VERSION AR068116.1 GI:5999338
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 3495)
AUTHORS Wilks/A.Frederick. and Harpur/A.
TITLE Protein tyrosine kinase
JOURNAL Patent: US 5852184-A 2 22-DEC-1998;
FEATURES Location/Qualifiers
source 1..3495
/organism="unknown"
BASE COUNT 1121 a 688 c 794 g 892 t
ORIGIN

Alignment Scores:
Pred. No.: 4.85e-05 Length: 3495
Score: 16.00 Matches: 16
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.46% Indels: 0
DB: Gaps: 0

US-09-397-967-16 (1-1099) x AR068116 (1-3495)

OY 1003 SeraspValTTPSerPheGlyValValLeuTyrGluLeuPheThrTyr 1018
|||||
Db 2695 TCAGATGTGTGAGCTTGGAGTGTCTCTATACGAACTTTTCACATAC 2742

RESULT 33
AR071176 AR071176 3495 bp DNA linear PAT 18-FEB-2000
LOCUS Sequence 2 from patent US 5910426.
ACCESSION AR071176
VERSION AR071176.1 GI:7222064
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 3495)
AUTHORS Wilks/A.Frederick., Ziemiecki/A. and Harpur/A.
TITLE Protein tyrosine kinase
JOURNAL Patent: US 5910426-A 2 08-JUN-1999;
FEATURES Location/Qualifiers
source 1..3495
/organism="unknown"
BASE COUNT 1121 a 688 c 794 g 892 t
ORIGIN

```

```

Alignment Scores:
Pred. No.: 4.85e-05 Length: 3495
Score: 16.00 Matches: 16
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.46% Indels: 0
DB: Gaps: 0

US-09-397-967-16 (1-1099) x AR071176 (1-3495)

OY 1003 SeraspValTTPSerPheGlyValValLeuTyrGluLeuPheThrTyr 1018
|||||
Db 2695 TCAGATGTGTGAGCTTGGAGTGTCTCTATACGAACTTTTCACATAC 2742

RESULT 34
AR177639 AR177639 3495 bp DNA linear PAT 17-DEC-2001
LOCUS Sequence 3 from patent US 6312941.
ACCESSION AR177639
VERSION AR177639.1 GI:17919994
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 3495)
AUTHORS Carter-Su,C., Rul,L.,Y. and Karow,D.S.
TITLE Compositions and methods for identifying signaling pathway agonists
and antagonists
JOURNAL Patent: US 6312941-A 3 06-NOV-2001;
FEATURES Location/Qualifiers
source 1..3495
/organism="unknown"
BASE COUNT 1121 a 688 c 794 g 892 t
ORIGIN

Alignment Scores:
Pred. No.: 4.85e-05 Length: 3495
Score: 16.00 Matches: 16
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.46% Indels: 0
DB: Gaps: 0

US-09-397-967-16 (1-1099) x AR177639 (1-3495)

OY 1003 SeraspValTTPSerPheGlyValValLeuTyrGluLeuPheThrTyr 1018
|||||
Db 2695 TCAGATGTGTGAGCTTGGAGTGTCTCTATACGAACTTTTCACATAC 2742

RESULT 35
I62349 I62349 3495 bp DNA linear PAT 07-OCT-1997
LOCUS Sequence 2 from patent US 5658791.
ACCESSION I62349
VERSION I62349.1 GI:2480297
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 3495)
AUTHORS Wilks/A.Frederick., Ziemiecki/A. and Harpur/A.
TITLE Antibodies which specifically bind to proteins having tyrosine
kinase activity, wherein said proteins have more than one tyrosine
kinase domain, and no SH2 domains
JOURNAL Patent: US 5658791-A 2 19-AUG-1997;
FEATURES Location/Qualifiers
source 1..3495
/organism="unknown"
BASE COUNT 1121 a 688 c 794 g 892 t
ORIGIN

Alignment Scores:
Pred. No.: 4.85e-05 Length: 3495

```

Score: 16.00 Matches: 16
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.468 Indels: 0
DB: 6 Gaps: 0

US-09-397-967-16 (1-1099) x 162349 (1-3495)

QY 1003 SerAspValTrpSerPheGlyValValLeuTyrgIuLeuPheThyr 1018
|||||
DB 2695 TCAGATGTGTGGAGCTTGTGAGCTTCTATACGAACCTTTCACATAC 2742

RESULT 36

LOCUS 187981 3495 bp DNA linear PAT 10-AUG-1998

DEFINITION Sequence 2 from patent US 5716818.

ACCESSION 187981

VERSION 187981.1 GI:3407921

KEYWORDS

SOURCE

ORGANISM

Unclassified.

REFERENCE 1 (bases 1 to 3495)

AUTHORS Wilks, A. Frederick., Ziemlecki, A. and Harpur, A.

TITLE Protein tyrosine kinase

JOURNAL Patent: US 5716818-A 2 10-FEB-1998;

FEATURES Location/Qualifiers

source

1..3495

/organism="unknown"

BASE COUNT 1121 a 688 c 794 g 892 t

ORIGIN

Alignment Scores:

Pred. No.: 4.85e-05 Length: 3495
Score: 16.00 Matches: 16
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.468 Indels: 0
DB: 6 Gaps: 0

US-09-397-967-16 (1-1099) x 187981 (1-3495)

QY 1003 SerAspValTrpSerPheGlyValValLeuTyrgIuLeuPheThyr 1018
|||||
DB 2695 TCAGATGTGTGGAGCTTGTGAGCTTCTATACGAACCTTTCACATAC 2742

RESULT 37

LOCUS AF001362

DEFINITION Homo sapiens Jak2 kinase (JAK2) mRNA, complete cds.

ACCESSION AF001362

VERSION AF001362.1 GI:3236321

KEYWORDS

SOURCE

ORGANISM

Homo sapiens.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 3500)

AUTHORS Datal, I., Arpala, E., Dadi, H., Kulkarin, S., Squire, J. and

Rolfman, C.M.

TITLE Cloning and characterization of the human homolog of mouse Jak2

JOURNAL Blood 91 (3), 844-851 (1998)

MEDLINE 98118198

PUBMED 9446644

REFERENCE 2 (bases 1 to 3500)

AUTHORS Rolfman, C.M.

TITLE Direct Submission

JOURNAL Submitted (26-APR-1997) Immunology, Hospital for Sick Children, 555

University Ave., Toronto, Ontario M5G 1X8, Canada

Location/Qualifiers

source

1..3500

/organism="Homo sapiens"

/db_xref="taxon:9606"

/Chromosome="9"
/map="9p23-24"
/cell_line="pre-B leukaemia cell line G2"
/cell_type="thymocytes, peripheral blood lymphocytes"
1..3500
/gene="JAK2"
1..3399
/gene="JAK2"
/codon_start=1
/product="jak2 kinase"
/protein_id="AAC23653.1"
/db_xref="GI:3236322"

/translation="MGNACTMTMEGTSSTSIYONGDISGNANSMQIDPVLOYLY
HSICKSEADYITFPSSGEVVAEICTASAKAGCTIPYTHNMFALSETERLYPNHYF
HIDSTRINVLTRIRREYPRMGSSNAYRHNGISRGAEVPLDDEVWSYLFQMRHD
EVHGMIKVPVTHETQEBGLPMVAVDMMRIARENDPLAIVNSISYKFLPKCRARI
QDYLILRKRIYRFRRIIOFQSOCKATARNILKLYLITETLDSAFTEFEVKEEG
SGPSGEIEFATIIITNGGIOMSRGKHSESELTEDQILQCDPNIIDVSIKOAND
GSNSRVYTIHKODGNLEIELSLREALSFVSLIDGYRLADAHYICKVEAPVAY
LENIOSNCHGPIISMDPAISKLKAGNOTGLVLRCSPDFKAYFLTAVAREVYIEK
HCLITKNNEPEYVNSGTRKNSSSLKDLNLCYOMETVRSNIIIFPTCKCPKPKDKS
LLVFRINVSVPPTSPILQPTFHMOYFHHIRNEDLIFNESLQGTFTKLEKVRRE
VGDYQGLHETEVLLKLVLDKAHINYSESEFSAASMSKLSHHLVNLNGVCGVCE
VQEFVKGSLDTYLLKKNKNCINIIMKLEVAQOLAMHFEENTLIGNVCAKNILIL
VEEDRTKGNPEFLKIDPGISITVLPKRIQDERIPWVPECIEPNKNLNATDPSFG
TTLMWICSGGDKPLSALDSORKLOFEYEDRHQLPAPKNAELANLINNCMDVDPKPSF
RAIRIDNSLFTPPYELITENDMLPNMRIGALGSCAFEDPDPTOFERHILKPTQQLC
KCNFSVEMCKVDPDLODNTGCVAVAKLOHSTEBHLRFEIEILSLQHDNIVKIK
GVCTSGAGRRNKLIMEYLPYGLSDYDQKHKERIDHILKLYTSQIDKGMGYDITKRY
IHRLATRNILIVENENRYKIDGFLITVLPDKKEYVYKVEGESPIFVWYAPESLIESK
FSVADWSFGVVLVELETTYIEKSKSPAEFRMIGNDKQCOMIIFHLIELKNNGRIL
PRDGCDEIYIMTECMNNNVNORPSRDLALRVQVDRDMAC"

BASE COUNT 1181 a 608 c 740 g 971 t

ORIGIN

Alignment Scores:

Pred. No.: 4.86e-05 Length: 3500
Score: 16.00 Matches: 16
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.468 Indels: 0
DB: 9 Gaps: 0

US-09-397-967-16 (1-1099) x AF001362 (1-3500)

QY 1003 SerAspValTrpSerPheGlyValValLeuTyrgIuLeuPheThyr 1018
|||||
DB 3103 TCAGATGTGTGGAGCTTGTGAGCTTCTATACGAACCTTTCACATAC 3150

RESULT 38

LOCUS AR136148

DEFINITION Sequence 8 from patent US 6136595.

ACCESSION AR136148

VERSION AR136148.1 GI:14476820

KEYWORDS

SOURCE

ORGANISM

Unknown.

REFERENCE 1 (bases 1 to 3629)

AUTHORS Ihle, J., Witthuhn, B.A. and Silvennoinen, O.

TITLE Jak kinases and regulations of cytokine signal transduction

JOURNAL Patent: US 6136595-A 8 24-OCT-2000;

Location/Qualifiers

source

1..3629

/organism="unknown"

BASE COUNT 1142 a 732 c 843 g 912 t

ORIGIN

Alignment Scores:

Pred. No.: 5e-05 Length: 3629
Score: 16.00 Matches: 16
Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.46% Indels: 0
DB: 6 Gaps: 0

US-09-397-967-16 (1-1099) x ARL36148 (1-3629)

OY 1003 SeraspValTrpSerPheGlyValValLeuTyrGluLeuPheThrTyr 1018
|||||
DB 3196 TCAGATGTGTGAGCTTTGGAGTGTCTCTATACGAACTTTTCACATAC 3243

RESULT 39

ARL43945 ARL43945 3629 bp DNA linear PAT 08-AUG-2001
LOCUS Sequence 8 from patent US 6210654.
DEFINITION ARL43945
ACCESSION ARL43945
VERSION ARL43945.1 GI:15105812
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 3629)
AUTHORS Ihle,J., Witthuhn,B.A., Quelle,F.W. and Silvennoinen,O.
TITLE Jak Kinases and Regulation of cytokine signal transduction
JOURNAL Patent: US 6210654-A 8 03-APR-2001;
FEATURES Location/Qualifiers
source 1..3629
/organism="unknown"

BASE COUNT 1142 a 732 c 843 g 912 t

Alignment Scores:
Pred. No.: 5e-05 Length: 3629
Score: 16.00 Matches: 16
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.46% Indels: 0
DB: 6 Gaps: 0

US-09-397-967-16 (1-1099) x ARL43945 (1-3629)

OY 1003 SeraspValTrpSerPheGlyValValLeuTyrGluLeuPheThrTyr 1018
|||||
DB 3196 TCAGATGTGTGAGCTTTGGAGTGTCTCTATACGAACTTTTCACATAC 3243

RESULT 40

192575 3629 bp DNA linear PAT 01-DEC-1998
LOCUS Sequence 8 from patent US 5728536.
DEFINITION 192575
ACCESSION 192575
VERSION 192575.1 GI:3937045
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 3629)
AUTHORS Ihle,J.N., Witthuhn,B.A., Quelle,F.W. and Silvennoinen,O.
TITLE Jak Kinases and Regulation of Cytokine signal transduction
JOURNAL Patent: US 5728536-A 8 17-MAR-1998;
FEATURES Location/Qualifiers
source 1..3629
/organism="unknown"

BASE COUNT 1142 a 732 c 843 g 912 t

Alignment Scores:
Pred. No.: 5e-05 Length: 3629
Score: 16.00 Matches: 16
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.46% Indels: 0
DB: 6 Gaps: 0

US-09-397-967-16 (1-1099) x 192575 (1-3629)

OY 1003 SeraspValTrpSerPheGlyValValLeuTyrGluLeuPheThrTyr 1018
|||||
DB 3196 TCAGATGTGTGAGCTTTGGAGTGTCTCTATACGAACTTTTCACATAC 3243

RESULT 41

MUSPTRK 3629 bp mRNA linear ROD 27-JAN-1997
LOCUS Mouse protein-tyrosine kinase (JAK2) mRNA, complete cds.
DEFINITION 116956
ACCESSION 116956
VERSION 116956.1 GI:309463
KEYWORDS protein-tyrosine kinase.
SOURCE Mus musculus.
ORGANISM Mus musculus.
REFERENCE 1 (bases 1 to 3629)
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognath; Muridae; Murinae; Mus.
and Ihle,J.N., Witthuhn,B.A., Quelle,F.W., Cleveland,J.L., Yi,T.
TITLE Structure of the murine Jak2 protein-tyrosine kinase and its role
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 90 (18), 8429-8433 (1993)
MEDLINE 93391367
PUBMED 8378315
FEATURES Location/Qualifiers
source 1..3629
/organism="Mus musculus"
/db_xref="taxon:10090"
1..3629
/gene="JAK2"
94..3483
/gene="JAK2"
/codon_start=1
/evidence=experimental
/product="protein-tyrosine kinase"
/protein_id="AAB41327.1"
/db_xref="GI:409584"
/translation="MGACITMTMEMASTSPYHONGDIPGSANSVQTEPVLQVLY
HSLQAGEYLPKPSGEYVAEICVAASKACGTPYVHNKFMALSERETIWPYNNYF
HIDESTRHDLIRYREFPHWYCGSSRTYRGVSRGAEPALDDPFWSLFVQWRD
FVHGWIWVPTVHTHTEOECILMAVLDMMRIKKEKQDPLAVNSYKTFLLPKVRAKI
ODYHILTRKIRYRFRFIOFSQCKATANKLKYLINLETLOSAYTQDFEYKESA
RGPSGEIEFATITITGNGGIOWSGRKESEETLEDOVOLCEPDIDVSIQAOKE
CSNRSRYVHAKDQGVLELELSIKKALFVSLIDGYRLTDADHYLCKEVAPEV
LENTSHCGHPIISMDPAISIKKAKNGTGLYVRCSPKDKKYEFLPAVRENVIEK
HCLITKNEBGLNISTGNRNSLKLKDLNCTQETLVASDSITIFOTKCCPKRDKSN
LVERTGISDVOISPTLORNNVNWVFEKIRNEDLIENESLGQGTFTLIRGVRE
VGDFGOLHTEVEVLKVDKARHNSSEFPAASMSQLSKRLVLANGVCGEENLE
VOEFVFGSLDITYLKKNSINILMKLGVAKKOLAMAHFLEKSLINGVCANILLI
REEDRTGNPPEIKLSDPGISITVLPDIOERIPVPPPECIENPKRLNATDKMSFG
TTIMETSGGDKPLSALDSORKIOFYEDKQOLAPKWTETLANLINNMQYEDPRPAF
RAVIRDLNLSLTPDYELLTENDMLPNRIGALGSCAFEDBDTPQPERILKFLQOLG
KGNFGSVMCKRYDPLQNTGEVAVAKRLQSTHEHLDFERELEILSLQHDNLYTK
GVCSACGRNRLRLIMEYLPYGLSLDYLOKHKERIDHKRLQYTSQIKQMEYLGTRY
IHRDLATRNILIVENENRVKIGDFGLTAVLPQDKRYVVKRPGESPIRWVAPQSLTESK
FSVADWSVSGCVLYELFTYIEKSPVEERMKMGDKOGOMIVFLIELLSNGRL
PRPGCDPEIIVITTEGMNWNVSGRPSFRDLSFQWIKGIV"

BASE COUNT 1142 a 732 c 842 g 913 t

Alignment Scores:
Pred. No.: 5e-05 Length: 3629
Score: 16.00 Matches: 16
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.46% Indels: 0
DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x MUSPTRK (1-3629)

OY 1003 SeraspValTrpSerPheGlyValValLeuTyrGluLeuPheThrTyr 1018
|||||
DB 3196 TCAGATGTGTGAGCTTTGGAGTGTCTCTATACGAACTTTTCACATAC 3243

RESULT 42
RNU13396 3731 bp mRNA linear ROD 26-OCT-1995
LOCUS Rattus norvegicus protein-tyrosine kinase (JAK2) mRNA, complete
DEFINITION cds.
ACCESSION U13396
VERSION U13396.1 GI:1040849
KEYWORDS
SOURCE .
ORGANISM rat.
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 3731)
AUTHORS Duhe, R.J., Rul, H., Greenwood, J.D., Garvey, K. and Farrar, W.L.
TITLE Cloning of the gene encoding rat JAK2, a protein tyrosine kinase
JOURNAL Gene 158 (2), 281-285 (1995)
MEDLINE 95331631
PUBMED 7607553
REFERENCE 2 (bases 1 to 3731)
AUTHORS Duhe, R.J.
TITLE Direct Submission
JOURNAL Submitted (12-AUG-1994) Roy J. Duhe, Program Resources,
Inc./Dyncorp., National Cancer Institute - Frederick Cancer
Research and Development Center, P.O. Box B, Frederick, MD
21702-1201, USA
COMMENT On Oct 26, 1995 this sequence version replaced gi:595682.
FEATURES
source
1..3731
location/Qualifiers
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/cell_line="Mb2-Sp"
/cell_type="pre-T lymphoma"
/note="Long-Evans parent strain"
1..3731
/gene="JAK2"
157..3555
/gene="JAK2"
/note="cytokine receptor-related"
/codon_start=1
/product="protein tyrosine kinase"
/protein_id="AA079911.1"
/db_xref="GI:595683"
/translation="MGMACLVITMEGTSTPAHNGDIPGANSYKQTEPYLVQVLY
HSLGAGEGYLAFNGEYAEELCYAASACGTPYHNMFLMSETERIMVPPNHF
HIDESTRNHILIRIRYFPHWYCSNRTARYGSGAEPALIDPVMISYLPQNRHD
FVHGMIKVPVTEHETOECLEGMVLDMMIRIAKEKDOTPLAVNSISYKPLPCVRKI
QDYLILTRKIRIRYRFRFIQOFSOCKATARNLIKYLINLETLOSAYFEQPEVESA
RQSGSEIFATITITINGGIGIOWSRGKHRESEITLEDOLYCDPPIIDIVSKANOE
CSRESYVYVHKODCKVEIEISLKEALSPSLIDGYRLTADAHVYCKEVPAPAV
LENHNSCHGPISMDFAISRLKAGQTLVLRCSPKDNKYFLFAVERENVLEYK
HCLITNENGEYNLSGTRKNSFLADNLNLCYOMETVRSDSIIFOFTKCCPPKDKSN
LIVFRTNGVSDVQLSPTLQRHNHVNQMVPHKIRNEDLINESLGGSTFKIRGVARE
VDYQGLHTEVLLKVLDAKHNYSSEFEASMSQSLSKHLIVLNYGVCVCGEENIL
VOEFYFGSLDITYLKNNKNSINILMKLGAVALAMAMHLEEKSLIHGVACAKNILLI
REDRRTGNPPTIKLSDPGISTIVLPKOLIORIPVAPRECLENKNTLATDKSPFG
TTLMECSGCKPPLSLDSQRLQLETDHQLPAPRWTELANLINTCMYEDEDFRAP
RAVIRDLNLSLPDVELLENDLPMWRKALGAFSGAFEDRPTQFEENHLLFLQLG
KNGFSEVEMCRYPDLODNGEVAAVKIKQHSIEHLDFEREIEILSKSOHNITAYK
GVYCSAGRNRLILMEYLPYGLSDLVQKHKERIDHKRLLOYTSQICKMEYLGTRK
IHRDLATRLIVENENRKYIGDGLTKVLPQDEYKVKVEGESPIEMWAPSLTESK
FSVASDVMFGVIVELFTYIEKRSKSPPEPRMNCINDGOMIYPHILIELLKNRRL
PRPEGCDITIVIMTEPCNNNNNRPSEFDSLRLVDQLRDSMA"

BASE COUNT 1149 a 763 c 881 g 936 t
ORIGIN

Alignment Scores:
Pred. No.: 5, 11e-05 Length: 3731
Score: 16.00 Matches: 16
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.46% Indels: 0

DB: 10 Gaps: 0
US-09-397-967-16 (1-1099) x RNU13396 (1-3731)
QY 1003 SerAspValTrpSerPheGlyValLeuTyrCluLeuPheThrTyr 1018
Db 3259 TCAGATGTGTGACGCTTTGGTGTGCTGTACGACACTTTTCAGCTAT 3306
RESULT 43
AF148993 3739 bp mRNA linear VRT 29-JUN-2000
LOCUS Cyprinus carpio janus kinase 3 mRNA, complete cds.
DEFINITION AF148993
ACCESSION AF148993
VERSION AF148993.1 GI:6690507
KEYWORDS
SOURCE Cyprinus carpio.
ORGANISM Cyprinus carpio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Cyprinus.
REFERENCE 1 (bases 1 to 3739)
AUTHORS Yin, Z. and Kwang, J.
TITLE Molecular cloning of a carp Jak3 from activated leukocytes
JOURNAL Vet. Immunol. Immunopathol. 74 (3-4), 225-236 (2000)
MEDLINE 10802290
PUBMED 10802290
REFERENCE 2 (bases 1 to 3739)
AUTHORS Yin, Z. and Kwang, J.
TITLE Direct Submission
JOURNAL Submitted (06-MAY-1999) Laboratory of Animal Health Biotechnology,
Institute of Molecular Agrobiolology, National University of
Singapore, 1 Research Link, Singapore 117604, Singapore
FEATURES
source
1..3739
location/Qualifiers
/organism="Cyprinus carpio"
/strain="Koi"
/db_xref="taxon:7962"
/cell_type="head kidney leukocytes"
/note="con A induced"
49..3429
/codon_start=1
/product="Janus kinase 3"
/protein_id="AA024169.1"
/db_xref="GI:6690508"
/translation="MSEEEVPLMISERAGSOKSNMDSALQVHLVYSPINSSTPTI
STRHTAENVCAIAKASGILPVFNELALASEDLSVYVPPHYKRSSENIKVRV
PEFSPPGSGASARYFSISRRICAVLDYCYITDLPKQSSDPSVSGSGISPALSLQ
QECLEMAVLDRKAKKERNQSLGECNTTSTKSCUPETHRODIOIRSLARQIKRTL
KFLKILKRCASVERILKLYLMEISLVEPNVGSFPLHHSGLMDEQSEODRVKAVRY
SGEGIIQITKESQEMQFCDFPQITDISIKRLSQEOMPLEDIONVYTGITSEFAVHL
FHTLLEALSFVSLVDGYFRLTDTDSYHSCAEVAPSLIEDIONVCHGITSEFAVHL
KKAAGKNMFLRLSPKDFQKFLVCGIOTHLGMYKCKLEKNKFLIAGIHNSFCS
LKQLPDFQHSITLMSDIPVTLKCCPPRPRLTMTILRNSMTEMSSPMLORHKP
SHMORHMKHEDTMSBSLGGSTPHIRFGSKTDORQCVYTTTEVLLVLANHNCM
ESFEPASALMSQISIKHLIVGISVHRSKNIMQEFVKAALDLYLKSTSVYSVMK
LDVAQKALCALMLEEKNIAVIGNICAKMLLAREDDPTDSNPFKIKSDPINNALIG
KDVIYADRIPWAAPEVLDLLELEECDDKWSFGTLMELIRNNECAPLOGODLHKOOFE
HESNLPALMELEADLILAOQMOYQELPSCRSILROLNSLITSDYETLHAGTLPKS
DGFMRPMFKROODQVEERLEIRFISLIGGNESVSELCVDPMDGTGLVAVKEL
OSNKATVTDQREIOTISLSLHCDIVYKIGYCSGTSLKRWMEYIPTYSLGTYME
KHROWVCRRLLLPASQICKMEYLOSRYVHRDAAANLIVASDTLAKINDGFTKI
IPVDEYIRVTPGSGPVFWAPESISELKSFSKSDVNSFVILHELPSYCDISNPO
KIVYEDRKLCEPVPMWLSIFLFSKDMWRLPAPQCCPPKHNIMQCMERNSEDRPS
FSSLQDLIENSLDERECCKGIRVTLNRSILYFQIMYT"

BASE COUNT 1045 a 844 c 888 g 962 t
ORIGIN

Alignment Scores:
Pred. No.: 5, 12e-05 Length: 3739
Score: 16.00 Matches: 16
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.46% Indels: 0

DB: 5 Gaps: 0

US-09-397-967-16 (1-1099) x AFL48993 (1-3739)

QY 824 leuGLYLVGSLVAspPhcGLySerValGluLeuGlyAspArgTyAspPro 839
|||||

Db 2527 CTTGGGAAGGGTAACTTCGACAGTGTTCAGCTTTGTCTGTTAAGATCCG 2574

RESULT 44

AB036337 4042 bp mRNA linear MAM 08-JAN-2000

LOCUS AB036337

DEFINITION Sus scrofa mRNA for kinase-defective JAK2 variant, complete cds.

ACCESSION AB036337

VERSION AB036337.1 GI:6635253

KEYWORDS kinase-defective JAK2 variant.

SOURCE Sus scrofa adult muscle thoracic aortic cell cDNA to mRNA, clone:50D1.

ORGANISM Sus scrofa

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.

AUTHORS Ito, Y., Minezawa, M. and Awata, T.

TITLE Sus scrofa mRNA for JAK2 alternative splicing variant, complete cds

REFERENCE Published Only in Database (1999)

2 (bases 1 to 4042)

1. 4042

Location/Qualifiers

1. 4042

/organism="Sus scrofa"

/db_xref="taxon:9823"

/clone="50D1"

/cell_type="thoracic aortic cell"

/tissue_type="muscle"

/dev_stage="adult"

1. 4042

/gene="JAK2"

407. .3697

/gene="JAK2"

/codon_start=1

/product="kinase-defective JAK2 variant"

/protein_id="BAA8632.1"

/db_xref="GI:6635254"

/translation="MGMACTITTEMEGTSSTPHVNGNDIPGNSVKKQIDPVLQVLYL
HSLGAEGLYLFPAGETVAEETICVAASACGITTIVYISMFLAMETERIYPPNHV
HYDESTRHNVLYRIFFPYWYCGNSNRYRGISRGAEAPLLDFVMSYLFPAQRHD
FLYGWVKIPVTHETDEECIGMAVLDMRIAKEDQTPLDIVSSVSKTFELPCVAKI
ODYHLTRKRIYRRFRRIEORSHCKATARNLKLKYLINLETLQSAFTKEFEVKEP
RGSGEELFATIIITNGGIGMSRGKHSEETLTDODLQYCDPDIIDVSKOANOE
GSNERIYTIHKODKSLIEIETLSIREALSVSLIDGYRLADAAHYLCKEVAAPV
LENIQSNCHGPTSMDFALSKLKAQNGTGLVLRSPDKFKYFLTRAVERNVYEVK
HCLITRNGEENYLSGTNRKNSFLKDLNLCYOMETVRSNIIIFQTKCCPPKPKDSN
LLVFRTNGISDVPTSLQRRNNVQWYFHKIRNEDLI.FNESLQGTFTKIFKGVARE
VGDYGLHETVELLVLDKARHNSSESEFEASMSKSHKHLVANGCYVCGEENIL
VOEYKFGSLDYLTKKKNKSNITLMLKLEVAKOLAMAMHLEKTLIHGNVCAKNILII
NEEDRTKGNPPFIKLSDPGISTITVLPKTILORIIPWPECELENKKNILATDKMSFG
TTLMETCSGDKPLNALSQRLQFYEDRHQLPAPKWTLANLINCMADYEDDPRPSF
RAIIRDLNLSLFTPDYELLTENDMLPMNRIGALGSGAEEDBDPTQFEERHLKFLQDG
KCNFGSVEMCRYPDLODNTGEVAVKLGQSTEEHILRDFEREIETLSKQMDNIYVK
GVCYAGRNRNLKIMEYLPYGLRDYLOKHKERIDHKIKLOYTSQICKGMEYLGTRKY
IHRDLATRNIIIVLENRNRVKGIDPGLTKVLPQDKKEYVKKYKPEGESPIFWAPESLTSK
ESVADSWMSFGVVLIELFTYIETKSKSPRAVSVLNFQLEFFPNMSKAPPKIKLARI
LEPHITTYL"

4042

/gene="JAK2"

BASE COUNT 1290 a 737 c 875 g 1140 t

ORIGIN

Alignment Scores:

Pred. No.: 5 45e-05 Length: 4042

Score: 16.00 Matches: 16

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 1.46% Indels: 0

DB: 4 Gaps: 0

US-09-397-967-16 (1-1099) x AB036337 (1-4042)

QY 1003 SeraspValTTPSerPhcGLyValValLeuTyrgLluLeuPheThrTYR 1018
|||||

Db 3509 TCACATGTTTGAGAGCTTGAGTGGTCTGTATGTAACATTTTCACATAT 3556

RESULT 45

AF005216 4161 bp mRNA linear PRI 30-OCT-1997

LOCUS AF005216

DEFINITION Homo sapiens receptor-associated tyrosine kinase (JAK2) mRNA,

ACCESSION AF005216

VERSION AF005216.1 GI:2570358

KEYWORDS complete cds.

SOURCE Homo sapiens.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

AUTHORS Peeters, P., Raynaud, S.D., Cools, J., Wlodarska, I., Grosgeorge, J., Philip, P., Monpoux, F., Van Rompaey, L., Baens, M., Van den Berghe, H. and Marynen, P.

TITLE Fusion of TEL, the ETS-variant gene 6 (ETV6), to the receptor-associated kinase JAK2 as a result of t(9;12) in a lymphoid and t(9;15;12) in a myeloid leukemia

2 (bases 1 to 4161)

1. 4161

Location/Qualifiers

1. 4161

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="9"

/map="9p24"

1. 4161

/gene="JAK2"

495. .3893

/gene="JAK2"

/note="Janus kinase 2"

/codon_start=1

/product="receptor-associated tyrosine kinase"

/protein_id="AAB82092.1"

/db_xref="GI:2570359"

/translation="MGMACTITTEMEGTSSTSIYONGDISGNANSMKQIDPVLQVLYL
HSLGSEADYLFPSGGEVAEETICVAASACGITTIVYISMFLAMETERIYPPNHV
HYDESTRHNVLYRIFFPYWYCGNSNRYRGISRGAEAPLLDFVMSYLFPAQRHD
FLYGWVKIPVTHETDEECIGMAVLDMRIAKEDQTPLDIVSSVSKTFELPCVAKI
ODYHLTRKRIYRRFRRIEORSHCKATARNLKLKYLINLETLQSAFTKEFEVKEP
RGSGEELFATIIITNGGIGMSRGKHSEETLTDODLQYCDPDIIDVSKOANOE
GSNERIYTIHKODKSLIEIETLSIREALSVSLIDGYRLADAAHYLCKEVAAPV
LENIQSNCHGPTSMDFALSKLKAQNGTGLVLRSPDKFKYFLTRAVERNVYEVK
HCLITRNGEENYLSGTNRKNSFLKDLNLCYOMETVRSNIIIFQTKCCPPKPKDSN
LLVFRTNGISDVPTSLQRRNNVQWYFHKIRNEDLI.FNESLQGTFTKIFKGVARE
VGDYGLHETVELLVLDKARHNSSESEFEASMSKSHKHLVANGCYVCGEENIL
VOEYKFGSLDYLTKKKNKSNITLMLKLEVAKOLAMAMHLEKTLIHGNVCAKNILII
NEEDRTKGNPPFIKLSDPGISTITVLPKTILORIIPWPECELENKKNILATDKMSFG
TTLMETCSGDKPLNALSQRLQFYEDRHQLPAPKWTLANLINCMADYEDDPRPSF
RAIIRDLNLSLFTPDYELLTENDMLPMNRIGALGSGAEEDBDPTQFEERHLKFLQDG
KCNFGSVEMCRYPDLODNTGEVAVKLGQSTEEHILRDFEREIETLSKQMDNIYVK
GVCYAGRNRNLKIMEYLPYGLRDYLOKHKERIDHKIKLOYTSQICKGMEYLGTRKY

polya_site
BASE COUNT 1491 a 810 c 954 g 1344 t
ORIGIN
RAIIRDLNSLFTPDYELTENDMLPMNRICALGSGAFEDRPTQFEERHKLFLQOLG
KAGVT
/gene="JAK2"
4599

Alignment Scores:
Pred. No.: 6.04e-05 Length: 4599
Score: 16.00 Matches: 16
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.46% Indels: 0
DB: 4 Gaps: 0

US-09-397-967-16 (1-1099) x AB036336 (1-4599)

QY 1003 SeraspValTrrpSerPheGlyValValLeuTyrGluLeuPheThrTyr 1018
|||||
Db 3206 TCAGATGTTTGAGCTTTGGAGTGTCTGTATGACTTTTCACATAT 3253

RESULT 48
AB006011 4812 bp mRNA linear MAM 29-JUL-1997
LOCUS Sus scrofa mRNA for JAK2, complete cds.
DEFINITION AB006011
ACCESSION AB006011 GI:2285795
VERSION JAK2.
KEYWORDS
SOURCE Sus scrofa
ORGANISM Sus scrofa adult muscle thoracic aortic cell cDNA to mRNA.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
1 (sites)
Ito, Y., Mikawa, S., Kobayashi, E., Wada, Y. and Minezawa, M.
Domestic pig mRNA for JAK2, complete cds
Unpublished
2 (bases 1 to 4812)
Ito, Y.
Direct Submission
Submitted (24-JUL-1997) Yoshiyasu Ito, National Institute of Animal
Industry, Animal Genome Research Team; 2 Ikenodai, Kikizaki-machi,
Inashiki-gun, Ibaragi 305, Japan (E-mail: yoshiito@niai.affrc.go.jp,
Tel:0298-38-8627)

FEATURES
source
Location/Qualifiers
1..4812
/organism="Sus scrofa"
/db_xref="taxon:9823"
/cell_type="thoracic aortic cell"
/tissue_type="muscle"
/dev_stage="adult"
317..3712
/codon_start=1
/product="JAK2"
/protein_id="AA021662.1"
/translation="MGMAQLTMEEMEGTSTSPVHONGDIPGNANSVKOIDPYLOVLYL
HSLGAEADYLFKPAEGEVAEIEICIAASKAGITPPVHNMFLMSETERIWPNNHVF
HIDESTRHNVLYRIRFYPWYCGSNRTYRHGISRGAEAPLDDEVMSTLFAQWRHD
FLYGWVKIPVTHEOEELGMAVLDMMRIKAKENDQPLDIYSSVSKFTLFGVRAKI
ODYHILTRKRIARYRFRFTIOFSHCKATARNIKLKYLINLETLOSATFTEKFEVKEG
SGPGEELFATIIITNGGCIOMSRGKRKESETLTEDOLXCDEPSNIDVSTIQANOE
GSNSRVTYTHKODKSLLEISLRELSLRLSVSLIDGYRILRTADAHYLCKEVAPVAV
LENIOSNCHGPIISMDFAIKSLKAGNOTGLVLRCSPPDENKYLFLTFVBERENVTEK
HCLITTKNENEGYNSGTRKNFSNLKDLNCOYMETVRSDDIIFQTKCCPPKPKDKN
LLVPTNGVSDVPTSPLOPRTNHNQVFNHRIINREDLIFNESLGGFTKLFKGVRE
VGDGOLAEHEVILKLVLDKARHNSSEFEFAASMSKLSHKHLVANGVCACENIL
VOEYVKGCSIDTYIKKKKNKINILMKLEVAOLAMAMHLEKTLIHGVCANILILI
REEDRTKGNPPFIKLSDPGISITVLPRILOERIPWVPECEIENKNILNATDKWSFG
TTLWELICGGDKPLNALSOKRLOFEDRHOLPAKWAELANLNNCMYDEPRPSP
RAIIRDLNSLFTPDYELTENDMLPMNRICALGSGAFEDRPTQFEERHKLFLQOLG
KAGVT
KGNFSGVEMCYRYPDLONTGEVAVKRLQHSSTEBHLDFFEREIILSLQHDNTVYK
GVCYSAGARNRLIMELYLPGLSDYLOKHKERIDHKLLQYTSQICKMEIGTKRY
IHRDLATRNILVENENRYKIGDFILTKVLPQKKEYKKKEGESIPFWVAPESLTSK

polya_site
BASE COUNT 1554 a 853 c 1016 g 1389 t
ORIGIN
FVASDVMSFGVLYLELYETIEKSPSPAEFRMIGNDKOGMIVFHEILKNNGRRL
4812
/note="17 a nucleotides"
853

Alignment Scores:
Pred. No.: 6.27e-05 Length: 4812
Score: 16.00 Matches: 16
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.46% Indels: 0
DB: 4 Gaps: 0

US-09-397-967-16 (1-1099) x AB006011 (1-4812)

QY 1003 SeraspValTrrpSerPheGlyValValLeuTyrGluLeuPheThrTyr 1018
|||||
Db 3419 TCAGATGTTTGAGCTTTGGAGTGTCTGTATGACTTTTCACATAT 3466

RESULT 49
AX354568 5117 bp DNA linear PAT 06-FEB-2002
LOCUS Sequence 15 from Patent WO0179355.
DEFINITION AX354568
ACCESSION AX354568
VERSION AX354568.1 GI:18619370
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1
Hancock, M.W. and Ozkaynak, E.
Roles of jak/stat family members in tolerance induction
Patent: WO 0179355-A 15 25-OCT-2001.
Millennium Pharmaceuticals, Inc. (US)
Location/Qualifiers
1..5117
/organism="Homo sapiens"
/db_xref="taxon:9606"
478..3876
/note="unnamed protein product"
/codon_start=1
/protein_id="CA022981.1"
/db_xref="GI:18619371"
/translation="MGMAQLTMEEMEGTSTSPVHONGDIPGNANSVKOIDPYLOVLYL
HSLGAEADYLFKPAEGEVAEIEICIAASKAGITPPVHNMFLMSETERIWPNNHVF
HIDESTRHNVLYRIRFYPWYCGSNRTYRHGISRGAEAPLDDEVMSTLFAQWRHD
FLYGWVKIPVTHEOEELGMAVLDMMRIKAKENDQPLDIYSSVSKFTLFGVRAKI
ODYHILTRKRIARYRFRFTIOFSHCKATARNIKLKYLINLETLOSATFTEKFEVKEG
SGPGEELFATIIITNGGCIOMSRGKRKESETLTEDOLXCDEPSNIDVSTIQANOE
GSNSRVTYTHKODKSLLEISLRELSLRLSVSLIDGYRILRTADAHYLCKEVAPVAV
LENIOSNCHGPIISMDFAIKSLKAGNOTGLVLRCSPPDENKYLFLTFVBERENVTEK
HCLITTKNENEGYNSGTRKNFSNLKDLNCOYMETVRSDDIIFQTKCCPPKPKDKN
LLVPTNGVSDVPTSPLOPRTNHNQVFNHRIINREDLIFNESLGGFTKLFKGVRE
VGDGOLAEHEVILKLVLDKARHNSSEFEFAASMSKLSHKHLVANGVCACENIL
VOEYVKGCSIDTYIKKKKNKINILMKLEVAOLAMAMHLEKTLIHGVCANILILI
REEDRTKGNPPFIKLSDPGISITVLPRILOERIPWVPECEIENKNILNATDKWSFG
TTLWELICGGDKPLNALSOKRLOFEDRHOLPAKWAELANLNNCMYDEPRPSP
RAIIRDLNSLFTPDYELTENDMLPMNRICALGSGAFEDRPTQFEERHKLFLQOLG
KAGVT
KGNFSGVEMCYRYPDLONTGEVAVKRLQHSSTEBHLDFFEREIILSLQHDNTVYK
GVCYSAGARNRLIMELYLPGLSDYLOKHKERIDHKLLQYTSQICKMEIGTKRY
IHRDLATRNILVENENRYKIGDFILTKVLPQKKEYKKKEGESIPFWVAPESLTSK

Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.46% Indels: 0
 DB: 6 Gaps: 0

US-09-397-967-16 (1-1099) x AX354568 (1-5117)

OY 1003 SerAspValTrpSerPheGlyValLeuTyrGluLeuPheThrTyr 1018
 DB 3580 TCAGATGTTTGAGCTTGGAGTGTCTGTATGAACTTTCACATAC 3627
 RESULT 50
 AF058925 5117 bp mRNA linear PRI 23-JUN-1998
 LOCUS Homo sapiens Jak2 kinase mRNA, complete cds.
 DEFINITION AF058925
 ACCESSION AF058925
 VERSION AF058925.1 GI:3068751
 KEYWORDS
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 5117)
 AUTHORS Saltzman, A., Stone, M., Franks, C., Searfoss, G., Munro, R., Jaye, M.,
 and Trushchenko, Y.
 TITLE Cloning and characterization of human Jak-2 kinase: high mRNA
 JOURNAL expression in immune cells and muscle tissue
 MEDLINE Blochem. Biophys. Res. Commun. 246 (3), 627-633 (1998)
 PUBMED 9618263
 REFERENCE 2 (bases 1 to 5117)
 AUTHORS Saltzman, A., Stone, M., Franks, C., Searfoss, G., Munro, R., Jaye, M.,
 and Trushchenko, Y.
 TITLE Direct Submission
 JOURNAL Submitted (10-APR-1998) Gene Medicine, Rhone-Poulenc Rorer, 500
 FEATURES Arcola Rd., Collegeville, PA 19426, USA
 SOURCE location/Qualifiers
 1..5117
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 1..2393
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="I.M.A.G.E Consortium ID 686684"
 478..3876
 /codon_start=1
 /product="Jak2 kinase"
 /protein_id="AAC23982.1"
 /db_xref="GI:3068752"
 /translation="MGMACLTMTMEGTSSTIYONGDISGNANSMKQIDPVLOYLYL
 HSLKSEADYITFPFSGEYVAEITIASKACGCTIPYHNMFALMSETERIYPPNHYF
 HIDESTRHNVLYRIRFYPPRYCGSNMAYRHGSRGAEPALDIDFVYSLEAQRHD
 FVHMGIWVPVTHETQEDCLGAVLDMMHIAENQOTPLATNSISYTFELKCRAKI
 ODVHILTRKIRIRYRFRFIOFSQCKATARKLKLYLINTLETSAPYTERFEVERG
 SGSPGELEIFATIIITNGGIOWSRGKRESELTLEODOLCTDSDSNIIIDVIRKANE
 GSNESRYVTHIKQDKNLEISLREASVSLIDCYRYLTAAHYHLYCEVAPAY
 LENTQSCNCHGISMDPAISKAKNGNOTGLVLCSPRODFKTYLTRAVENYETK
 HCLITKNENEYVNSGTRKNSSSLKDLINCYOMETVRSNDIIFQFTCCPREDKSN
 LIAVPTNGVSDVPTSPTRQPTNNOMVFNHRTINEDLIFNSLQGGFTKLFKCVRE
 VQDYGQHLSETLEVLLVLDKKAHNRSESEFEASMSKLSHKLVLQNGCVCGENIL
 VOEFVFGSLDTYLVKKNKNCINILMKLEVAOKLAWAMHLEENLLIGNVCAKILLI
 REEDRKNGPPEITLSDPGISITVLPDIIOERTIPWPRECIENKUNLNTDWSRG
 TTLWEITSGGDKPLSALDSOKLQFTEDRHLPAKMAELANLINCMDEPDPRSP
 RAILRLDNLSTLPDYELLTENDMLPNMRIGALGSGAFEDNDPQFEERHLKLOQLS
 KGNFSGVMCRKYLQDNTGEVAVKRLQHSTEDHLDFEREIILSLQDNIVKTK
 GVCYSAGRNLKLIMEYLPYSLADYLOKHREIDHILKLOYTQICKGMXYLCTKRY
 IHRDLATRNILVENENRVIKIGDFGLTKVLPQDKERYVKEGESPIRWVAPESLTSK
 FSVASDWSYGVVLYELFTYTERKSPAPERMICNDKOCQIVPHILILKNNKRL
 PRPDCPEDEITMTCEGMNNVNRPSFRDLALRVDOIRDMWAG"

BASE COUNT 1633 a 927 c 1103 g 1464 t

Alignment Scores: 6.58e-05 Length: 5117

Score: 16.00 Matches: 16
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.46% Indels: 0
 DB: 9 Gaps: 0

US-09-397-967-16 (1-1099) x AF058925 (1-5117)

OY 1003 SerAspValTrpSerPheGlyValLeuTyrGluLeuPheThrTyr 1018
 DB 3580 TCAGATGTTTGAGCTTGGAGTGTCTGTATGAACTTTCACATAC 3627
 RESULT 51
 AC119228/C
 LOCUS Mus musculus clone RP24-209J17, LOW-PASS SEQUENCE SAMPLING.
 DEFINITION AC119228 66974 bp DNA linear HTG 25-APR-2002
 AC119228.1 GI:20303952
 ACCESSION AC119228
 VERSION AC119228.1 GI:20303952
 HTG: HTGS_PHASE0.
 KEYWORDS HTG: HTGS_PHASE0.
 SOURCE Mus musculus.
 ORGANISM Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 66974)
 AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N.,
 Anderson, S., Barin, N., Baslet, V., Bloom, T., Boguslavsky, L.,
 Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J.,
 Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A.,
 Cook, A., Cooke, P., Dearellano, K., Dewar, J., Diaz, J.S., Dodge, S.,
 Faro, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Garyna, S.,
 Glinde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
 Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
 Kamat, A., Karatas, A., Kells, C., Larocque, K., Lamazares, R.,
 Landers, T., Lehoczy, J., Levine, R., Lindblad-Toh, K., Liu, G.,
 Maclean, C., Macdonald, P., Major, J., Marcus, N., Matthews, C.,
 McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Meneus, L.,
 Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R.,
 Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D.,
 Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
 Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
 Roman, J., Roselli, M., Roy, A., Santos, R., Schauer, S., Schnupack, R.,
 Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
 Straus, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
 Topham, K., Travers, M., Travis, N., Trigglio, J., Vassiliev, H.,
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.-J., Young, G.,
 Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
 TITLE Direct Submission
 JOURNAL Submitted (25-APR-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 COMMENT All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html
 ----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L25230
 Center clone name: 209_J_17

* NOTE: This record contains 83 individual
 * sequencing reads that have not been assembled into
 * contigs. Runs of N are used to separate the reads
 * and the order in which they appear is completely
 * arbitrary. Low-pass sequence sampling is useful for
 * identifying clones that may be gene-rich and allows
 * overlap relationships among clones to be deduced.
 * However, it should not be assumed that this clone

* Will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved. 709: contig of 709 bp in length
* 1
* 710 809: gap of 100 bp
* 810 1526: contig of 717 bp in length
* 1527 1626: gap of 100 bp
* 1627 2342: contig of 716 bp in length
* 2343 2442: gap of 100 bp
* 2443 3133: contig of 691 bp in length
* 3134 3233: gap of 100 bp
* 3234 4033: gap of 100 bp
* 3934 4033: contig of 700 bp in length
* 4034 4730: contig of 697 bp in length
* 4731 4830: gap of 100 bp
* 4831 5519: contig of 689 bp in length
* 5520 5619: gap of 100 bp
* 5620 6322: contig of 703 bp in length
* 6323 6422: gap of 100 bp
* 6423 7140: contig of 718 bp in length
* 7141 7240: gap of 100 bp
* 7241 7964: contig of 724 bp in length
* 7965 8064: gap of 100 bp
* 8065 8793: contig of 729 bp in length
* 8794 8893: gap of 100 bp
* 8894 9622: contig of 729 bp in length
* 9623 9722: gap of 100 bp
* 9723 10458: contig of 736 bp in length
* 10459 10558: gap of 100 bp
* 10559 11247: contig of 689 bp in length
* 11248 11347: gap of 100 bp
* 11348 12047: contig of 700 bp in length
* 12048 12147: gap of 100 bp
* 12148 12842: contig of 695 bp in length
* 12843 12942: gap of 100 bp
* 12943 13635: contig of 694 bp in length
* 13637 13735: gap of 100 bp
* 13737 14466: contig of 730 bp in length
* 14467 14566: gap of 100 bp
* 14567 15276: contig of 710 bp in length
* 15277 15376: gap of 100 bp
* 15377 16098: contig of 722 bp in length
* 16099 16198: gap of 100 bp
* 16199 16924: contig of 726 bp in length
* 16925 17024: gap of 100 bp
* 17025 17760: contig of 736 bp in length
* 17761 17860: gap of 100 bp
* 17861 18550: contig of 690 bp in length
* 18551 18650: gap of 100 bp
* 18651 19353: contig of 703 bp in length
* 19354 19453: gap of 100 bp
* 19454 20158: contig of 705 bp in length
* 20159 20258: gap of 100 bp
* 20259 20973: contig of 715 bp in length
* 20974 21073: gap of 100 bp
* 21074 21780: contig of 707 bp in length
* 21781 21880: gap of 100 bp
* 21881 22600: contig of 720 bp in length
* 22601 22700: gap of 100 bp
* 22701 23397: contig of 697 bp in length
* 23398 23497: gap of 100 bp
* 23498 24216: contig of 719 bp in length
* 24217 24316: gap of 100 bp
* 24317 25052: contig of 736 bp in length
* 25053 25152: gap of 100 bp
* 25153 25888: contig of 736 bp in length
* 25889 25988: gap of 100 bp
* 25989 26692: contig of 704 bp in length
* 26693 26792: gap of 100 bp
* 26793 27478: contig of 686 bp in length
* 27479 27578: gap of 100 bp
* 27579 28290: contig of 712 bp in length
* 28291 28390: gap of 100 bp

* 28391 29098: contig of 708 bp in length
* 29099 29198: gap of 100 bp
* 29199 29903: contig of 705 bp in length
* 29904 30003: gap of 100 bp
* 30004 30709: contig of 706 bp in length
* 30710 30809: gap of 100 bp
* 30810 31535: contig of 726 bp in length
* 31536 31635: gap of 100 bp
* 31636 32329: contig of 694 bp in length
* 32330 32429: gap of 100 bp
* 32430 33137: contig of 708 bp in length
* 33138 33237: gap of 100 bp
* 33238 33969: contig of 732 bp in length
* 33970 34069: gap of 100 bp
* 34070 34760: contig of 691 bp in length
* 34761 34860: gap of 100 bp
* 34861 35566: contig of 706 bp in length
* 35567 35666: gap of 100 bp
* 35667 36371: contig of 705 bp in length
* 36372 36471: gap of 100 bp
* 36472 37166: contig of 695 bp in length
* 37167 37266: gap of 100 bp
* 37267 37982: contig of 716 bp in length
* 37983 38082: gap of 100 bp
* 38083 38784: contig of 702 bp in length
* 38785 38884: gap of 100 bp
* 38885 39591: contig of 707 bp in length
* 39592 39691: gap of 100 bp
* 39692 40413: contig of 722 bp in length
* 40414 40513: gap of 100 bp
* 40514 41246: contig of 733 bp in length
* 41247 41346: gap of 100 bp
* 41347 42078: contig of 732 bp in length
* 42079 42178: gap of 100 bp
* 42179 42874: contig of 696 bp in length
* 42875 42974: gap of 100 bp
* 42975 43677: contig of 703 bp in length
* 43678 43777: gap of 100 bp
* 43778 44472: contig of 695 bp in length
* 44473 44572: gap of 100 bp
* 44573 45279: contig of 707 bp in length
* 45280 45379: gap of 100 bp
* 45380 46053: contig of 674 bp in length
* 46054 46153: gap of 100 bp
* 46154 46860: contig of 707 bp in length
* 46861 46960: gap of 100 bp
* 46961 47646: contig of 686 bp in length
* 47647 47746: gap of 100 bp
* 47747 48458: contig of 712 bp in length
* 48459 48558: gap of 100 bp
* 48559 49262: contig of 704 bp in length
* 49263 49362: gap of 100 bp
* 49363 50090: contig of 728 bp in length
* 50091 50190: gap of 100 bp
* 50191 50902: contig of 712 bp in length
* 50903 51002: gap of 100 bp
* 51003 51740: contig of 738 bp in length
* 51741 51840: gap of 100 bp
* 51841 52537: contig of 697 bp in length
* 52538 52637: gap of 100 bp
* 52638 53332: contig of 695 bp in length
* 53333 53432: gap of 100 bp
* 53433 54094: contig of 662 bp in length
* 54095 54194: gap of 100 bp
* 54195 54877: contig of 683 bp in length
* 54878 54977: gap of 100 bp
* 54978 55686: contig of 709 bp in length
* 55687 55786: gap of 100 bp

Alignment Scores:
Pred. No.: 0.000515 Length: 66974
Score: 16.00 Matches: 16
Percent Similarity: 100.00% Conservative: 0


```

LENIQSNCHPISMDRAISKIKKAGNOTGIVLRCSKPDNKKEITFAVRENVIEYK
HCLITKNEEYVNLSTGKNPSSLKDLNCTQMETVASNDITFOFTKCCPKKXDSN
LVFERTNGVSDVPTLQRPTRHNNQVFEKIRNEDIIFNESLGQGTFTLIFGVRE
VGDYGOJHTEVLTAVLDKARHNSSESEFPAASMXSLSHKHVLANGVCGDENTL
VOEEFKGSLDPTVKKKKNKNCINILMKLEVAOKAMAHFLEENTLHIGNVCANILIT
FEEDRKGNPPEFIKLSDDPGISITVLPDIOERIPWVPECTEKNPLNATOKWSG
TTITMETCSGDKPLSALDSORRQOFYDRHOLAPKAEFLANINNCMDPEPRSEF
RAIIRDNLSLFTPDYELLTENDLPMNRIGALGSGAFEDRPTOFEEERLKLQOIG
KNGFSVEMCRDPLQDNTGEVAVKRLQSHTEHEHLEFERELEILKSLQHDNIVTK
GVCSAGRNKLKIMETLPGYSLRDYLOKHEERIDHKLQVTSQICGKGLQTKRY
IHRDLATRNILIVENENRVLIGDEGLTVLQDKEYYKVGESPIFWVYPESTLTSK
FSVADVMSFGVLYELFTYIEKSPAEFMRALIGDKOGMVFLLILBLKNKGL
PRDGCDEIYVIMTECMNNVNRPEFRLLALRVDDIRDNMAG"

repeat_region
/feature="LIM44A repeat: matches 5811. .6155 of consensus"
complement(5181. .5643)
/feature="match: GSS: Em:AQ128841"
6716. .6807

repeat_region
/feature="46 copies 2 mer tt 66% conserved"
7989. .8028

repeat_region
/feature="20 copies 2 mer tg 80% conserved"
8104. .8313

repeat_region
/feature="Charlize repeat: matches 3530. .3762 of consensus"
8494. .10889

repeat_region
/feature="TIGGER1 repeat: matches 1. .2416 of consensus"
complement(112366. .12931)
/feature="match: GSS: Em:AQ417363"
22026. .23724

misc_feature
/feature="JAK2"
/feature="Cpg Island"
/feature="not_experimental"
22236. .23231

gene
/feature="ba39K24.2"
/feature="pseudo"
22236. .23231

CDS
/feature="ba39K24.2"
/feature="ba39K24.2 (casein kinase pseudogene)"
match: proteins: Tr:O9VEX2 Sw:062761 Sw:062762 Sw:062763
Sw:P48730 Sw:P78368 Tr:O9X303 Sw:O9YGM4 Tr:O9VGE3
Sw:P35509 Sw:O06486 Sw:P40233 Sw:P70065 Sw:P40234
Tr:O82321 Sw:P05986 Tr:Q24302"
/pseudo
/codon_start=1
/feature="not_experimental"
24089. .24561

misc_feature
/feature="JAK2"
/feature="Cpg Island"
/feature="not_experimental"
complement(29564. .30075)
/feature="match: GSS: Em:B55561"
complement(32591. .32919)
/feature="match: GSS: Em:AQ067163"
complement(32615. .32882)
/feature="match: GSS: Em:AQ066533"
complement(32616. .33009)
/feature="match: GSS: Em:AQ340673"
33384. .34018

misc_feature
/feature="JAK2"
/feature="match: GSS: Em:AQ347464"
complement(35412. .36133)
/feature="match: GSS: Em:AQ096478"
35778. .35845

repeat_region
/feature="34 copies 2 mer tt 67% conserved"
39499. .39855

repeat_region
/feature="THEIC repeat: matches 1. .369 of consensus"
41257. .43368

repeat_region
/feature="TIGGER1 repeat: matches 265. .2417 of consensus"
43369. .43607

repeat_region
/feature="MER8 repeat: matches 1. .239 of consensus"
43608. .43876

repeat_region
/feature="TIGGER1 repeat: matches 1. .265 of consensus"
complement(51924. .52384)
/feature="match: GSS: Em:B68961"

```

```

misc_feature complement(51942. .52378)
/feature="match: GSS: Em:AQ230025"
51995. .52071

repeat_region
/feature="TIGGER2 repeat: matches 1. .77 of consensus"
52558. .52871

misc_feature
/feature="JAK2"
/feature="match: GSS: Em:AQ074500"
55574. .57969

repeat_region
/feature="TIGGER1 repeat: matches 1. .2417 of consensus"
complement(59285. .59816)
/feature="match: GSS: Em:AQ377376"
63327. .64290

repeat_region
/feature="LTP5 repeat: matches 1. .969 of consensus"
complement(66506. .66624)
/feature="match: GSS: Em:AQ390658"
complement(66507. .66624)
/feature="match: GSS: Em:AQ718621"
complement(66617. .66697)

misc_feature
misc_feature
misc_feature

Alignment Scores:
Pred. No.: 0.00109 Length: 171146
Score: 16.00 Matches: 16
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.46% Indels: 0
DB: 9 Gaps: 0

US-09-397-967-16 (1-1099) x ALI61450 (1-171146)

OY 1003 SeraspValITRPserPheGlyValValLeuTyrgIueuPheRrTYR 1018
|||||
DB 104338 TCAGATGTTTGAGCTTGGAGTGGTCTGTATGACATTTTCACATAC 104385

RESULT 54
AF173032 4744 bp mRNA linear ROD 15-AUG-1999
LOCUS AF173032 Mus musculus tyrosine kinase TYK2 (Tyk2) mRNA, complete cds.
DEFINITION AF173032
ACCESSION AF173032.1 GI:5733094
VERSION AF173032.1
KEYWORDS Mus musculus.
SOURCE Mus musculus.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murine; Mus.
REFERENCE
AUTHORS McCoy,B., Range,A., Lassnig,C., Karaghiousoff,M. and Muller,M.
TITLE Cloning and characterization of mouse Tyk2 cDNA
JOURNAL Unpublished
AUTHORS McCoy,B., Range,A., Lassnig,C., Karaghiousoff,M. and Muller,M.
TITLE Direct Submission
JOURNAL Submitted (28-JUL-1999) Animal Production and Genetics, University
of Veterinary Medicine, Veterinärplatz 1, Vienna 1210, Austria
FEATURES
source 1. .4744
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="8"
1. .4744
/feature="Tyk2"
216. .3758
/feature="Tyk2"
/codon_start=1
/product="tyrosine kinase TYK2"
/protein_id="AAd9423.1"
/db_xref="GI:5733095"
/translation="MPLGRRRAILEDKADTEAPLPTGCLMVLHMPDPEGGEPM
VTFPSOTSLAEVNCIHIAHKVGIPTPCINTEPALYNAQAKVLPNHLIDTSODMNLTF
RMRFYFRMHGMNPOEPVAVYRGFGPGSETSDRAQGVOLDSAFYELFEQKHEFM
NDVYSLRDISEEELIHFKNSLSIMAPLHICHLASGVPLEEMAREISFNKCIPISEF
ROHTRQHNVLTRLRHVRFRFLAARFGHLISQVAVYKVIATIERLAPRGSSRIY
CHLEVAQPERDPCYTIQNSGQTADPGELPSGPTHEVLTGTGGIOMPLQOESSE
KNSRGNHNGSRSGKKPKAPRAGELHLESPEPWTYFCDFDISHVYLKERVHILH

```

DNKCLLCLCSQAEALSFVALVDGYFRLTADSSHYLCHENAPRLVTSIONGHIPL
MDPEOAKIMPEGLYLQWSTSHLRLLTYAHNRPNAXSNRGLRLRKRPEITQPG
AFVLDGMSFSLGDLALOGCSLRAGDDCFPLHXCCLPPREISNLVIMRGRRAH
TRPLNLSOLFNRHODITOLSHLGSGCTRTNVRGILLRVGPDGKNDGCPPEPG
TSCQDLRLVLYLVDPSHHIDIALAFEXASLSQSVSHMLATLHGVCSGSEITVTEF
VEHGLDWLRKQVOPMTKRVVAQDLASLSTLEKRLVHGVNCGRNILLARLGL
EGTNPFTKLSDPGVGOALSHREERVERIPMTAVAPLSCGTSLSLTADMMFGATLL
EICFDGEAPLQGRGPEKERYTKKHQLPESSPELATLRTROCLTYEPADMRSPFTIL
RDLRLQONLVTGSAVNSDSPADSPVFNHKKRYLKKIDLGSGHGKVSILCYDPTND
GTGEVAVKALKEGGCPOLRSQOREITLRLTYEHKRYKKGCCCEDEGKSVQVME
VPLGSLDLYLPRHCVGAQLILFNOQCEGMAYLHACHYIHRDLAARNVLLDNRLV
KIDGGLAKAVPEGHEHYRVREDSDSPVFWAPECLKECKRYVASDVGFTVTELL
TYCDNSQSPHMKFTLIGTQGMVTLRTLELGERLPRDRCPCEIYHLMKNCWE
TEASFRPTFONLVPILOTAEKYCGVPSVFSVC"

BASE COUNT 1036 a 1358 c 1341 g 1006 t 3 others

ORIGIN

Alignment Scores:

Pred. No.: 0.000584 Length: 4744

Score: 15.00 Matches: 15

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 1.36% Indels: 0

DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x AF173032 (1-4744)

QY 333 GUAAlaLeuSerPheValAlaLeuValAspGlyTyrPheArgLeu 347
|||||
DB 1434 GAGGCCCTCTCTTGTGGCCCTGTCGTCATTCCTGCTTG 1478

RESULT 55

AF052607 7134 bp DNA linear ROD 05-SEP-1998

LOCUS Mus musculus non-receptor tyrosine kinase (Tyk2) gene, partial cds.

DEFINITION AF052607

VERSION AF052607.1 GI:3549252

KEYWORDS

SOURCE Mus musculus.

ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 7134)
Laasnig,C., Karaghiosoff,M., Range,A., Brem,G. and Muller,M.
Submitted (05-MAR-1998) Department of Veterinary Biotechnology and
Molecular Genetics, Institute of Animal Breeding and Genetics,
Veterinary University of Vienna, Veterinaerplatz 1, Vienna 1210,
Austria

FEATURES

SOURCE

Location/Qualifiers

1. 7134

/organism="Mus musculus"

/strain="129/Sv"

/db_xref="taxon:10090"

/chromosome="8"

/rname="Tyk2"

/rfeature="Tyk2"

/gene="Tyk2"

/note="member of Jak (Janus) family; contains Jak homology
regions 7 to 4"

/codon_start=3

/product="non-receptor tyrosine kinase"

/protein_id="AAC34380.1"

/db_xref="GI:3549253"

/translation="ITPPCLNLFALYNAAQKVMVLPNHLDTSDNNLTYRMRPYFN
WIKMNPQEPAYVRCGFGAETSSDRAGVQLDLSAFETLFEQGRKHFXNDVYSLRD

CDS

LSSEFIHFKNESLGMFLFLHIALTSRGVLEPMAREISFKNCIPRSFROHROH
VLTJRLRVRRLRRLARRRPHLSQVYWKVYLTLETLARFSSERLIPVCHLELAD
PERDPTIONSGQIADGPELPSGPPTHEVLVGTGIGIOMHPQIOSESNGNSGNP
HGSRSGRKPKAPEGEHLTESPOEPMTYFCDPDISHVVLKERVHTHLDNCLLL
CLCSQAEALSFVALVDGYFRLTADSSHYLCHENAPRLVTSIONGHIPL"

BASE COUNT 1684 a 1824 c 1807 g 1785 t 34 others

ORIGIN

Alignment Scores:

Pred. No.: 0.00081 Length: 7134

Score: 15.00 Matches: 15

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 1.36% Indels: 0

DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x AF052607 (1-7134)

QY 333 GUAAlaLeuSerPheValAlaLeuValAspGlyTyrPheArgLeu 347
|||||
DB 6939 GAGGCCCTCTCTTGTGGCCCTGTCGTCATTCCTGCTTG 6983

RESULT 56

AF090382 14907 bp DNA linear VRT 05-OCT-2000

LOCUS Tetracodon fluvialtilis jak2 tyrosine kinase (JAK2) gene, complete
cfs.

DEFINITION AF090382

VERSION AF090382.1 GI:5918519

KEYWORDS

SOURCE Tetracodon fluvialtilis.

ORGANISM Tetracodon fluvialtilis.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorphia; Acanthopterygii; Percormorpha; Tetraodontiformes;
Tetraodontidae; Tetraodon.
1 (bases 1 to 14907)
Lau,J.H., Yan,S.J., Lee,T.F., Chou,C.M., Chen,S.T., Hwang,P.P.,
Chou,C.K. and Huang,C.J.
Complete genomic organization and promoter analysis of the
round-spotted pufferfish JAK1, JAK2, JAK3, and TYK2 genes
DNA Cell Biol. 19 (7), 431-446 (2000)

REFERENCE

JOURNAL MEDLINE 20399330

PUBMED 10945233

AUTHORS

1 (bases 1 to 14907)

2 (bases 1 to 14907)

Leu,J.H. and Huang,C.J.
Submitted (03-SEP-1998) Biological Chemistry, Academia Sinica,
Taipei P.O. Box 23-106, Taiwan

FEATURES

SOURCE

Location/Qualifiers

1. 14907

/organism="Tetracodon fluvialtilis"

/db_xref="taxon:47145"

/gene="JAK2"

/rname="JAK2"

/rfeature="JAK2"

/gene="JAK2"

/note="member of Jak (Janus) family; contains Jak homology
regions 7 to 4"

/codon_start=3

/product="non-receptor tyrosine kinase"

/protein_id="AAC34380.1"

/db_xref="GI:3549253"

/translation="ITPPCLNLFALYNAAQKVMVLPNHLDTSDNNLTYRMRPYFN
WIKMNPQEPAYVRCGFGAETSSDRAGVQLDLSAFETLFEQGRKHFXNDVYSLRD

CDS

/product="ak2 tyrosine kinase"
 /protein_id="AA056051.1"
 /db_xref="GI:5918520"
 /translation="MASLAMEILADACTYONONTAORESEPEARLSTVDYLRLHFTYS
 GQGTQDSSFLSTYPPNGVAEELCIDAKKCSISLPTLSLFGFLRRDRDLMSFPHNTHQ
 VNGEMASDFVEFLRTRYEPFGYSGASRAARYAVAGSSSPVDFDCWSMFLFSQMRNF
 LDGMPKTPSSHELOEGLCAALADMTKREQLSPDLDEHTMAYKSFLEPELKAQO
 CNFLNTPAARLPFQOLAPKQSPGSHRAPTELLIMESLEAKAFYEFVPELSELR
 GOLLIVANGDLOKOEKPKQSDDELTOICDPVPVNTISIKASGEAPREAVTIN
 KQKSNLVEALSSEALSPYSLIDGYSRLTYRTHDANHLKEVAPRVLRYITSHCGPIS
 LERATINOLQCKGKQGLFILOCSPPKNDYFLTPFVYLYGVYERHCQITRSECGQPS
 LSGTRNFSSLLHEELSCYKREVRSDSVNFCQCCPKPAERSLCLCRNKCPENP
 LSPSTNHNISQVFKIRKEDLFENFMSAGOTFLPIKGVAKELGIDGLMHQTEVYM
 KYLDQARNRSEEPFKAAMQSLSHVHLILNYGVCQSEENIMVAGKSGSDTLV
 KKNSSNITMLKLEEVAKOLAAWNPFLKEEKNILPANGVACAKNVLILREDMRAGNPFK
 LDGSGSITVLPKEILVEYRPPWPCVNDPNLSTADKNSGFTGMEILSGEGRFL
 ASDLNKTKTLFYEDRDLQAPKKILELANITSCMDYEPFLPPTPALIIRDLHSITPDD
 LELTVSDSLIPNRTAGSTGSGFSDQEPQAEERHLLFLQOLKGNFGVSECRIDVP
 LDQNTGVAVVAKLQGHGTHLIDFERELTELKSLQDHNIIVLYGVCYSASRNKRLRV
 MEYLPQSDSLITLKKERIDHNKILVAGSSQCKMEYLSKRYIHRLAARNLVS
 ESRYAKIDFGKTLIPDKCYVYRPGSPFLVAPESLISEKRSYAVDSWGSVVL
 YELTSSNNSRSPPYTMSMGNDKGGOLLVHLIELLKSGSRLPQPLDCTPIHEITMEIN
 EOCDDNPNRYSPESEKELALSIDLFRHREK"

Alignment Scores:	
Pred. No.:	0.00146
Score:	15.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	1.36%
BB:	5
Length:	14907
Matches:	15
Conservative:	0
Mismatches:	0
Indels:	0
Gaps:	0

US-09-397-967-16 (1-1099) x AF090382 (1-14907)

Qy	1003	SeraspValTrrPserPheGlyValValLeuTyrgIuLeuphetrr	1017
Db	12204	TCAGATCTCTGGAGCTTCGGGGTGGTGTCTATGAGCTCTTCAAC	12248

RESULT	57
AC027154/c	
LOCUS	172936 bp DNA linear
DEFINITION	Mus musculus chromosome 6 clone RP23-37F10 map 6, WORKING DRAFT
ACCESSION	AC027154
VERSION	AC027154.3 GI:12621447
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE	Mus musculus.

REFERENCE	AUTHORS	TITLE	JOURNAL	REFERENCE
1 (bases 1 to 172936)	Birren, B., Linton, L., Nusbaum, C. and Lander, E.	Mus musculus chromosome 6, clone RP23-37F10	Unpublished	2 (bases 1 to 172936)
	Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,			

Boguslavsky, L., Boukhalter, B., Brown, A., Burkett, G.,
Campodiano, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S.,
Collamore, A., Cooke, P., Deavallano, K., Dewar, R., Diaz, J. S.,
Dodge, S., Domingo, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D.,
Galagan, J., Gardyna, S., Ginde, S., Coyette, M., Graham, L.,
Grand-pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L.,
Howard, J. C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,
Klein, J., Lacroque, K., Lamazares, R., Landers, T., Lechoczky, J.,
Levine, R., Lien, C., Liu, G., Locke, K., MacDonald, P., Marquis, N.,
McCarthy, M., Meenan, P., McGuck, A., McKenna, K., McPheters, R.,
Melidim, J., Menus, L., Mihova, T., Miranda, C., Mlena, V., Morrow, J.,
Naylor, T., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P.,

TITLE	Direct Submissions
JOURNAL	Submitted (28 MAR-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT	On Jan 31, 2001 this sequence version replaced g119367491.
	Clinical.C., Pollara,T.M., Oliver,J., Peterson,K., Pierre,N., Pisanil,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D., Roy,A., Santos,R., Schanover,S., Severy,P., Spencer,B., Stange-Hopmann,N., Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W., Young,G., Zainnov,J., Zimmer,A. and Zody,M.

Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>
 --- genome center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIMR
 Web site: <http://www-seq.wi.mit.edu>
 Contact: sequence_submissions@genome.wi.mit.edu
 --- Project Information
 Center project name: L5743
 Center clone name: 37_F_10
 --- Summary Statistics

Sequencing vector: M13; M7615; 5% of reads
Sequencing vector: Pml3d; n/a; 5% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 170054 bases at least Q40
Consensus quality: 171641 bases at least Q30
Consensus quality: 172109 bases at least Q20
Insert size: 15890; agarose-fp
Insert size: 12536; agarose-ctrl
Quality coverage: 8.3 In Q20 bases; agarose-fp
Quality coverage: 9.5 In Q20 ba.

* NOTE: This is a 'working draft' sequence. It currently
* consists of 5 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of 'N', but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 8836: contig of 8836 bp in length
* 8837 gap of 100 bp
* 8927 93243: contig of 8407 bp in length
* 9324 gap of 100 bp
* 9334 110724: contig of 17881 bp in length
* 11075 110824: gap of 100 bp
* 110825 167746: contig of 36524 bp in length
* 167749 167848: gap of 100 bp
* 167849 1729336: contig of 5088 bp in length

FEATURES	Location/Qualifiers
source	1. .172936

misc_feature

misc_feature

misc_feature

Miscellaneous

BASE COUNT 44231 a 43573 c 42293 g 42434 t 405 others
ORIGIN
Alignment Scores:
Pred. No.: 0.0104 Length: 172936
Score: 15.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.36% Indels: 0
DB: 2 Gaps: 0
US-09-397-967-16 (1-1099) x AC027154 (1-172936)
OY 333 GUAUAlaUSeRpheValAlaUSeValaSpGlyTYPheArgLeu 347
|||||
Db 128207 GAGGCCCTGCTTGTGGCCCTGTCGATGCTATTTCGCTTG 128163
RESULT 58
LOCUS AF091238 8788 bp DNA linear VRT 05-OCT-2000
DEFINITION Tetraodon fluvialtilis JAK3 tyrosine kinase (JAK3) gene, complete
ACCESSION AF091238
VERSION AF091238.1 GI:5739117
KEYWORDS Tetraodon fluvialtilis.
SOURCE Tetraodon fluvialtilis.
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodon.
1 (bases 1 to 8788)
Leu,J.H., Yan,S.J., Lee,T.F., Chou,C.M., Chen,S.T., Hwang,P.P.,
Chou,C.K. and Huang,C.J.
Complete genomic organization and promoter analysis of the
round-spotted pufferfish JAK1, JAK2, JAK3, and TYK2 genes
DNA Cell Biol. 19 (7), 431-446 (2000)
20399330
PUBMED 10945233
2 (bases 1 to 8788)
Yan,S.J. and Huang,C.J.
Direct Submision
Submitted (09-SEP-1998) Institute of Biological Chemistry, Academia
Sinica, PO Box 23-106, Taipei, Taiwan, Republic of China
FEATURES
source
1..8788
/organism="Tetraodon fluvialtilis"
/db_xref="taxon:47145"
2163..8636
/gene="JAK3"
join(2163..2449,2652..2913,2988..3105,3170..3287,3378..3523,3596..3899,3984..4103,4196..4353,4492..4606,4710..4890,5039..5163,5321..5449,5633..5717,5821..5945,6075..6213,6328..6479,6618..6762,6917..7044,7123..7312,7387..7511,7610..7782,7993..8110,8194..8307,8384..8503)
/gene="JAK3"
/product="JAK3 tyrosine kinase"
join(2715..2913,2988..3105,3170..3287,3378..3523,3596..3899,3984..4103,4196..4353,4492..4606,4710..5039,5163,5321..5449,5633..5717,5821..5945,6075,6213,6328..6479,6618..6762,6917..7044,7123..7312,7387..7511,7610..7782,7993..8110,8194..8307,8384..8503)
/gene="JAK3"
/codon_start=1
/product="JAK3 tyrosine kinase"
/protein_id="A050352.1"
/db_xref="GI:5739118"
/translation="MNIAEEFAPLVIRDRGSGTSSTSSAGSLQVHLVFPATKAT
TTHMSTGQIAENVICQAGKSGILPVYIGFSLASADISFWPSPHMTENLHVC
FRVFFGNNFDGPRACRYSLTRDTSPVLDHSLTAVLFLNRLTRDFTSSGSLSP
LTQDDICGLAVDLWRLEKQSSVYKDLCKSVSYSCPLSHRVIKRNWVERRI
RNLKRLFKRFGSCAVDDCNLKLMLTLAGLQPSVGSSEFHHVDPSSSLNSAVSL
VQVTEIGTISSECEGVWQFCEKIDTISIKRICREOVNDSRMVITRKDDAC

LEVAPOSKEALSLVSDYGFRLTDTSTHYFCODTAPPSILLEGIKNHQPIINSEFA
VNKLKRWSEDEGTFLLRKSPDYGEFFLTVAAGTPLGLDKDLVILKNNHSLPVS
KFSFLKELFYVCOONKLLADVAPVLSRCCPPSPSOETLNLIIRNSFPVLAOLSP
EGNKLHIHFOFOMIYEDLOMGESHGOSFTLFPKYSKSNISEGKRVFLELDVNH
RNLMSFEFASLMSOFSHKHLILVYCATLAKNIMQOEPEVYALDYLKRGSVSV
SMKLDVAKOASLYTFLEONIVGNICAKULLARQSDPSQSLPPPIKISDPISVL
MLGADVDRITPMYAPVLAPELVLSPEMLTLESDKWSGATLWELFNNGNPLLGMDLMAF
YESQQLPESQWTELDLINOCDYQADFRSCRSITROLFLITSDYVILFATEPT
QEPAMRALTPSOQDOTVEERHRLYISLILGNGSVLELCCYDPLGNGTELVAVKL
QPSKSSMEDFRKEVETLSVLCIEYIKVGVCSMGRLSMGLVTEVLPYSGILGYE
NNRHKVDTRMLLEASOICKMEYLQSLRFRHDLARNILVASLSVKTADGTRIT
IPCCKEYRYTQPSGSPITFWAPPSINSESRSHSDVSGVVALTELSYCDINSNK
RLYQOIGHVQTPESISLHLANILKSNMRLPAPPDCAKVISLMECEMLDFMERPCF
SFLNOLLEVIMODERDSFKG.
BASE COUNT 2291 a 2068 c 2109 g 2320 t
ORIGIN
Alignment Scores:
Pred. No.: 0.00903 Length: 8788
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.27% Indels: 0
DB: 5 Gaps: 0
US-09-397-967-16 (1-1099) x AF091238 (1-8788)
OY 1003 SerAPVALTPSerPheGlyValValLeuTyrGluLeuPhe 1016
|||||
Db 8036 TCCGACGTTTGAGTGTGGAGTGCTTTACGAGCTCTTC 8077
RESULT 59
LOCUS AR002620 41 bp DNA linear PAT 04-DEC-1998
DEFINITION Sequence 9 from patent US 5741899.
ACCESSION AR002620
VERSION AR002620.1 GI:3964174
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 41)
AUTHORS Capon,D.J., Tjian,H., Smith,D.H., Winslow,G.A. and Siekevitz,M.
TITLE ChimERIC receptors comprising Janus kinase for regulating cellular
pro liferation
JOURNAL Patent: US 5741899-A 9 21-APR-1998;
FEATURES
source
1..41
/organism="unknown"
BASE COUNT 7 a 9 c 15 g 10 t
ORIGIN
Alignment Scores:
Pred. No.: 0.00116 Length: 41
Score: 13.00 Matches: 13
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.18% Indels: 0
DB: 6 Gaps: 0
US-09-397-967-16 (1-1099) x AR002620 (1-41)
OY 1 MetAlaProSeRglugIurhrProLeuIleProGln 13
|||||
Db 40 ATGGCACCTCCAGTGAGGAGACACCTCGATCCCTCAG 2
RESULT 60
LOCUS AR099670 41 bp DNA linear PAT 14-FEB-2001
DEFINITION Sequence 9 from patent US 6077947.
ACCESSION AR099670
VERSION AR099670.1 GI:12809436
KEYWORDS

SOURCE	Unknown.
ORGANISM	Unclassified.
REFERENCE	1 (bases 1 to 41)
AUTHORS	Capon,D.J., Tian,H., Smith,D.H., Winslow,G.A. and Siekevitz,M.
TITLE	DNA encoding an intracellular chimeric receptor comprising Janus kinase
JOURNAL	Patent: US 6077947-A 9-20-JUN-2000;
FEATURES	Location/Qualifiers
source	1..41
BASE COUNT	/organism="unknown"
ORIGIN	7 a 9 c 15 g 10 t
Alignment Scores:	
Pred. No.:	0.00116 Length: 41
Score:	13.00 Matches: 13
Percent Similarity:	100.00% Conservative: 0
Best Local Similarity:	100.00% Mismatches: 0
Query Match:	1.18% Indels: 0
DB:	Gaps: 6
US-09-397-967-16 (1-1099) x AR099670 (1-41)	
OY	1 MetalaProSerGluThrProLeuIleProGln 13
Db	40 ATGGACCTCCAGTGAGGACACTCTGCATCCTCAG 2
RESULT 61	
LOCUS	AR002619 45 bp DNA linear PAT 04-DEC-1998
DEFINITION	Sequence 8 from patent US 5741899.
ACCESSION	AR002619
VERSION	AR002619.1 GI:3964173
KEYWORDS	.
SOURCE	Unknown.
ORGANISM	Unknown.
REFERENCE	Unclassified.
AUTHORS	1 (bases 1 to 45)
TITLE	Capon,D.J., Tian,H., Smith,D.H., Winslow,G.A. and Siekevitz,M. Chimeric receptors comprising janus kinase for regulating cellular proliferation
JOURNAL	Patent: US 5741899-A 8-21-APR-1998;
FEATURES	Location/Qualifiers
source	1..45
BASE COUNT	/organism="unknown"
ORIGIN	11 a 16 c 10 g 8 t
Alignment Scores:	
Pred. No.:	0.00125 Length: 45
Score:	13.00 Matches: 13
Percent Similarity:	100.00% Conservative: 0
Best Local Similarity:	100.00% Mismatches: 0
Query Match:	1.18% Indels: 0
DB:	Gaps: 6
US-09-397-967-16 (1-1099) x AR002619 (1-45)	
OY	1 MetalaProSerGluThrProLeuIleProGln 13
Db	6 ATGGACCTCCAGTGAGGACACTCTGCATCCTCAG 44
RESULT 62	
LOCUS	AR099669 45 bp DNA linear PAT 14-FEB-2001
DEFINITION	Sequence 8 from patent US 6077947.
ACCESSION	AR099669
VERSION	AR099669.1 GI:12809435
KEYWORDS	.
SOURCE	Unknown.
ORGANISM	Unknown.
REFERENCE	Unclassified.

1 (bases 1 to 45)
Capron,D.J., Tien,H., Smith,D.H., Winslow,G.A. and Siekevitz,M.
DNA encoding an intracellular chimeric receptor comprising Janus
kinase
Patent: US 6077947-A 8 20-JUN-2000;
Location/Qualifiers
1. 45
/organism="unknown"
BASE COUNT 11 a 16 c 10 g 8 t
ORIGIN

Alignment Scores:
Pred. No.: 0.00125 Length: 45
Score: 13.00 Matches: 13
Percent Similarity: 100.00% Conservative: 0
Best local Similarity: 100.00% Mismatches: 0
Query Match: 1.18% Indels: 0
DB: 6 Gaps: 0

US-09-397-967-16 (1-1099) x AR099669 (1-45)

QY 1 MetalapropioSerGluGluThrPheProLeuIleProgin 13
|||||
Db 6 ATGCACCTCCAGAGAGAGACACCTCTGATCCCTCAG 44

RESULT 63
LOCUS AY078944 1592 bp mRNA linear PLN 09-MAR-2002
DEFINITION Arabidopsis thaliana At2g28940/F8N16.23 mRNA, complete sequence.
ACCESSION AY078944
VERSION AY078944.1 GI:19310416
KEYWORDS FLI CDNA.
SOURCE Arabidopsis thaliana.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 1592)
Shinn,P., Chen,H., Cheuk,R., Kim,C.J., Meyers,M.C., Banh,J.,
Bowser,L., Carninci,P., Chang,E., Dale,J.M., Goldsmith,A.D.,
Hayashizaki,T., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G.,
Kawai,J., Lam,B., Lee,D.M., Lin,J., Miranda,M., Narusaka,M.,
Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T.,
Satou,M., Seki,M., Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C.,
Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W.,
Theologis,A. and Ecker,J.R.
Arabidopsis CDNA clones
Unpublished
2 (bases 1 to 1592)
Shinn,P., Chen,H., Cheuk,R., Kim,C.J., Meyers,M.C., Banh,J.,
Bowser,L., Carninci,P., Chang,E., Dale,J.M., Goldsmith,A.D.,
Hayashizaki,T., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G.,
Kawai,J., Lam,B., Lee,D.M., Lin,J., Miranda,M., Narusaka,M.,
Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T.,
Satou,M., Seki,M., Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C.,
Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W.,
Theologis,A. and Ecker,J.R.
Direct Submission
Submitted (08-FEB-2002) Salk Institute Genomic Analysis Laboratory
(SIGNAL), Plant Biology Laboratory, The Salk Institute for
Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037
USA
There is no suitable ORF in this CDNA clone.
The Salk, Stanford, PGEC (SSP) Consortium members carried out the
sequencing and annotation of the RAFL cDNAs: Shinn,P., Chen,H.,
Cheuk,R., Kim,C.J., Meyers,M.C., Banh,J., Bowser,L., Chang,E.,
Dale,J.M., Goldsmith,A.D., Jones,T., Karlin-Neumann,G., Lam,B.,
Hayashizaki,Y. and Shinozaki,K.

Lee, J.M., Lin, J., Miranda, M., Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Southwick, A., Tang, C.C., Toriumi, M., Wu, H.C., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Davis, R.W., Theologis, A., and Ecker, J.R.

Shim, P. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinzaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk) contributed equally to this work as PIs.

Location/Qualifiers
1. 1592

FEATURES
source
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/chromosome="2"
/clone="RAF109-69-D18(R21702)"
/note="ecotype: Columbia"

BASE COUNT
ORIGIN
462 a 270 c 403 g 457 t

Alignment Scores:

Pred. No.: 0.0217 Length: 1592
Score: 13.00 Matches: 13
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.18% Indels: 0
DB: 8 Gaps: 0

US-09-397-967-16 (1-1099) x AY078944 (1-1592)

Qy 1003 SerAspValTrpSerPheGlyValIleuTyGluLeu 1015
|||||
Db 948 AGTGATGTATGACCTTGGGAGTGTGTATGAACCTG 986
|||||

RESULT 64

TFU53213 16977 bp DNA linear VRT 03-FEB-1998
LOCUS
DEFINITION
Tetraodon fluvialilis JAK1 tyrosine kinase (JAK1) gene, complete cds.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

U53213
U53213.1 GI:2828795
Tetraodon fluvialilis.
Tetraodon fluvialilis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoclestei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodon.

REFERENCE
AUTHORS
TITLE
1 (bases 1 to 16977)
Leu, J.H., Chang, M.S., Yao, C.W., Chou, C.K., Chen, S.T. and Huang, C.J.
Genomic organization and characterization of the promoter region of
the round-spotted pufferfish (Tetraodon fluvialilis) JAK1 kinase
gene

JOURNAL
MEDLINE
PUBMED
9434151
9434151

REFERENCE
AUTHORS
TITLE
JOURNAL
2 (bases 1 to 16977)
Leu, J.H., Chang, M.S., Chang, G.D. and Huang, C.J.
Direct Submision
Submitted (01-APR-1996) J.H. Leu, Inst. Biological Chemistry,
Academia Sinica, P.O. Box 23-106, Taipei, Taiwan 106, Republic of
China

FEATURES
source
Location/Qualifiers
1. 16977

gene
/organism="Tetraodon fluvialilis"
/db_xref="taxon:47145"
/clone="pr11"
5236. 16216
/gene="JAK1"
Join(5236..5440,6078..6201,7605..7758,7987..8138,
8245..8513,8714..8823,8930..9115,9233..9390,9592..9706,
9836..10034,11673..11782,11920..12072,12201..12288,
12367..12494,12615..12753,13464..13615,13691..13844,
14295..14383,14525..14717,14786..14910,15297..15469,
15563..15680,15876..15986,16109..>16216)

CDS

/gene="JAK1"
Join(5236..5440,6078..6201,7605..7758,7987..8138,
8245..8513,8714..8823,8930..9115,9233..9390,9592..9706,
9836..10034,11673..11782,11920..12072,12201..12288,
12367..12494,12615..12753,13464..13615,13691..13844,
14295..14383,14525..14717,14786..14910,15297..15469,
15563..15680,15876..15986,16109..16216)

/note="protein tyrosine kinase"
/codon_start=1
/product="JAK1 tyrosine kinase"
/protein_id="AAC00207.1"
/db_xref="GI:2828796"

/translation="MSTQVWELSMQCGKRRGRKAHSSPPSPFGLIEHTYTY
VHQLTSGCTTAEELVEARKCSISPLCNLFALDEKNNMYPNVEETVEELS
LKLHYRMFEYFRNMHNGTEGSSVWRHICSLKRGSLSPQKIPVGTPLDASLELYF
OGQSHFNKGLVPRKSHSDTDOHEIENECLEMAVALITHTNVGAPSHDISYKRFIP
ETLNRSIKORSFLPIRINNYEKKFSEFHORTVRSNITPYDLKIKYATLEGTSG
LGSEVFEPISLIVAOEGELGNGYVRYFNQGOAVKVSROTVYVGTGISMK
KPEMALTSKDKSKSKSKADGKQOONRKRKRETCMEYFCDFEITTAIVIKDTVTI
VRODNMGELOMASRAESLRYALVDGFRITVDHHYLCRDVAPASVNRINISCHG
PISTDEYAKHLKRGNEGYILWMSCTDFEYIVITVCTEELDKESRPVQYKNFOI
EESSDGFRLYGTDCRPTLMLEHLSQSRLTDNLOFOLRVCPQPPREISNLVYT
KDRVLTSPSPDQSLSPHRLKEDIQEEHIGRGTNTIYSGTLRVKSEEDAGYS
SFQEVKVLKLSYGHNDISLAFFETASMMRVSHKHLVILYGVCAHONIMVEEYV
QLGPLEVEMRROGSLSPMKFOVALASLSTYLEDKRLVHGVCACKNILLADGCG
TDEGGPFTKLSDPGIPITVLTRECVRIPIAECYKNSALSVADKMGFTTME
ICYGEVPLKPKKITEQKERYETECOLATPDCELLELHMCNVPDKRPFERALV
RDIWMLSEONPSIKPKPAEVDVPEKRGKRVKRDGEGHFGVEICARPDKDGK
ELVAVKSLKPEQDOSTNLSEIDIKALFHEINIVYKRGKIGCEGGAQRLMEYVP
MGSLLKDYLPRIHKDPTSLATILSYVQICGMEYIGSNVYIHRDLAARVAVENRBMK
IGDGLTKRSIKDNEGYTYVDKENDSPVEMVAPELTCCKRYFLASDWSFQVTELYLT
YDSSKSPMTCFLDMIGMTQOMVNLVNLNDGRLLPRPDGCPREYVELMKRCHEQ
KPDRIETFTKIKELSMQOEPQ"

BASE COUNT
ORIGIN
4600 a 3563 c 3734 g 3080 t

Alignment Scores:

Pred. No.: 0.144 Length: 16977
Score: 13.00 Matches: 13
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.18% Indels: 0
DB: 5 Gaps: 0

US-09-397-967-16 (1-1099) x TFU53213 (1-16977)

Qy 335 LeuSerPheValAlaLeuValAspGlyTrpPheArgLeu 347
|||||
Db 9263 CTGTCCTTGTAGCTTGTGATGGGTATTCGCGCTG 9301
|||||

RESULT 65
AC005727/c 107097 bp DNA linear PLN 11-MAR-2002
LOCUS
DEFINITION
Arabidopsis thaliana chromosome 2 clone FBNI6 map m154, complete
sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AC005727
AC005727.3 GI:20197435
HTG.
Arabidopsis thaliana.
Arabidopsis thaliana.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

REFERENCE
AUTHORS
TITLE
JOURNAL
2 (bases 1 to 107097)
Shen, M., Rouning, C.M., Fraser, C.M., Somerville, C.R. and Venter, J.C.
Unpublished
Lin, X.

REFERENCE
AUTHORS
TITLE
JOURNAL
Direct Submission
Submitted (09-MAR-2000) The Institute for Genomic Research, 9712
Medical Center Dr., Rockville, MD 20850, USA

```
REFERENCE 3 (bases 1 to 107097)
AUTHORS Town,C.D. and Kaul,S.
TITLE Direct Submission
JOURNAL Submitted (27-FEB-2002) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA, cdtown@tigr.org
COMMENT On Apr 18, 2002 this sequence version replaced gi:5598485.
FEATURES
    source
        1..107097
            /organism="Arabidopsis thaliana"
            /cultivar="Columbia"
            /db_xref="taxon:3702"
            /chromosome="2"
            /map="m154"
            /clone="F8N16"
            1..3005
            /note="overlap with BAC clone T1P11 (:14595..17599)."
```

```
misc-feature
    187..258
        /gene="At2g28730"
        187..258
            /gene="At2g28730"
            /product="tRNA-Cys"
            /complement(678..1179)
            /gene="At2g28740"
            /note="F8N16.2; identical to GB:M17133 and
            GB:M17132; supported by full length cDNA: Ceres:32642"
            complement(<678..>1179)
            /gene="At2g28740"
            complement(827..1138)
            /gene="At2g28740"
            /codon_start=1
            /product="histone H4"
            /protein_id="AAC79580.1"
            /db_xref="GI:3927823"
            /translation="MSGKGGKGLGKGAKRHKVLRDNIQTGTPAIRLRARGGV
            KRISGLIYEETRGVLKIFLENVITDQATYIEHARKKVTAMDVYVALKRGGRLLYGGG
            G"
            complement(1191..1211)
            /rpt_family="AT_rich"
            1543..1590
            /rpt_family="AT_rich"
            2811..3958
            /gene="At2g28750"
            /note="F8N16.3; pseudogene, reverse transcriptase"
            /pseudo
            complement(5356..5643)
            /gene="At2g28755"
            /note="F8N16.4"
            complement(join(<5356..5421,5539..>5643))
            /gene="At2g28755"
            complement(join(5356..5421,5539..5643))
            /gene="At2g28755"
            /codon_start=1
            /product="putative GNDP-glucose 4-6-dehydratase"
            /protein_id="AAM15077.1"
            /db_xref="GI:20197439"
            /translation="WEGIKRLMAGDKSGPINIGNGEFSIYELAEYKALIKPDEIK
            IVENIPDDPQRS"
            6797..6822
            /rpt_family="AT_rich"
            complement(7653..9826)
            /gene="At2g28760"
            /note="F8N16.5"
            complement(join(<7653..7835,7913..8017,8271..8337,
            8478..8571,8665..8742,8834..8987,9083..9174,9261..9435,
            9743..>9826))
            /gene="At2g28760"
            complement(join(7653..7835,7913..8017,8271..8337,
            8478..8571,8665..8742,8834..8987,9083..9174,9261..9435,
            9743..9826))
            /codon_start=1
            /product="putative nucleotide-sugar dehydratase"
            /protein_id="AAC79582.1"
```

```
repeat_region
    /rpt_family="GAAA)n"
    10935..11006
        /gene="At2g28770"
        10935..11006
            /gene="At2g28770"
            /product="tRNA-His"
            complement(11283..14459)
            /gene="At2g28780"
            /note="F8N16.7; predicted by genscan and genefinder"
            complement(join(<11283..12968,13077..13244,13929..>14459))
            /gene="At2g28780"
            complement(11283..12968,13077..13244,13929..14459))
            /codon_start=1
            /product="hypothetical protein"
            /protein_id="AAC79583.1"
            /db_xref="GI:3927826"
            /translation="MTERGRAMWRCTASAFRTALACTIVGSAITLYGPEWIRNVAFP
            AFSYVILITITPDLGDTLRGCMALVATGCSVGAIVTLKLRPRLTAETLALAA
            ALAAFYVILPNSSTHLVAKRIALGOVLIVIGIYIGATDVMHQLQVASTALGV
            ACVALIIVPLPRTATGCEVSKCEKLEONATYTRIKYKAFCDSDSMSAVASQARVL
            ARSSKLTIVTKRKYOSMTMERPEKIMNOVNDKRGKLSOMETALGMYAVASK
            SPILSSULGAEVKEDIKNIOERYTISIKVYNNSSQSVYPSDPPKPDCELTGLEIP
            GTPQDLPFYFEFLCIRLETLIIAKPEKNAKYLEKETRNSWISMDSKTIPMAKLT
            SLSGLAIIIGSMFSPNGYWAQIPVAFPAAREKTEVTRNKAAGTIGTVYGMG
            CEVQKELVYRFLSLPLPFLSSFLRSKMYGQAGGISALIAEGLILGKNGPSEF
            AIERIETETGLSCSINVELVQPTAAINAKIEISRSFHAECASLEGAASAKADI
            MESOKLRSHLNEIKKETADEAHAPSEFSPFSCYEKLFKLSKMDLLOFSGYAI
            GLCEGCKTKRSPCKCKEILSNVDKDLSTESITGLKSFEEITLKLSDALKALAKS
            DNTSWDIELKTNPSFSFSAVSPPEKILFTYLOHCSVADGLFRVEDEGEFEVYKS
            EVVLSLCAUGCEYERIGKETRELEENYKEVQSENSSHVNLHEISCKIRSLYK"
            complement(13246..13298)
            /rpt_family="AT_rich"
            15206..15240
            /rpt_family="AT_rich"
            complement(16081..16226)
            /rpt_family="(TATA)n"
            complement(17388..17489)
            /rpt_family="AT_rich"
            complement(17827..17848)
            /rpt_family="AT_rich"
            complement(18719..18749)
            /rpt_family="AT_rich"
            21595..21648
            /rpt_family="AT_rich"
            22793..22907
            /rpt_family="AT_rich"
            complement(23101..23146)
            /rpt_family="AT_rich"
            complement(23479..23502)
            /rpt_family="AT_rich"
            complement(23679..23701)
            /rpt_family="AT_rich"
            complement(25689..26616)
            /gene="At2g28790"
            /note="F8N16.8; supported by full length cDNA: Ceres:2058"
            complement(<25689..>26616)
            /gene="At2g28790"
            complement(25848..26597)
            /gene="At2g28790"
            /codon_start=1
            /product="putative thaumatin"
            /protein_id="AAC79584.1"
            /db_xref="GI:3927827"
```

```

/translation="MAKTSILPLAASELLISFSSAVDTSLFTLVNCPETVPALIO
PNAGHIVLEKGFALPTFTSRFNVPFTTHSGSRIMRTMCABHNGKESCLTGDGRL
ECNIGCGAPASLAOFDLHGHGHDPSGVYSIYDQYNDPMPTTPREGHCYCVYCR
EDLITKCPAHLORSHSGHYACKSCCEAFHDELCRCRHYNSPNTCKSHSLFFKH
ACPSSTFPAHDSFLMHDCASPRELVIRICH"
complement(27729..30415)
/gene="At2g28800"
/feature="F8N16.9: identical to GB:U89272: supported by full
length cDNA: Ceres:31457"
complement(join<27729..28078,28179..28328,28637..28740,
28833..28940,29082..29186,29288..29361,29446..29506,
29584..29638,29769..29879,29966..>30415))
/gene="At2g28800"
complement(join(27853..28078,28179..28328,28637..28740,
28833..28940,29082..29186,29288..29361,29446..29506,
29584..29638,29769..29879,29966..30342))
/gene="At2g28800"
/codon_start=1
/product="chloroplast membrane protein (ALBINO3)"
/protein_id="AAC79585.1"
/db_xref="GI:3927828"
/translation="MARLVSSPSSEFFGAPLTKPSSSLRHSVGSGGCTAOPLYRSNN
NKLFSTVRSFSLNEIPPHGLDSSVDIGAJFTTRASLLTYTADAAGVAGSVYTD
SSAVKSGMFGFISDAMELVILKIDGLSAVHVPAYGFAITLLITIKATYPTIK
OOVESTLAMONLOPKIKAQCRAGNDEIQLTSLIKQAGVNPAGLCPLTATIPV
WIGLYALSNVANEGLFTEGFMTIPSLGSPITSAKQSGSISWLFEPYDGHPLPMY
DTVAIVLVPLIASQVSMETKPPQDPAOKNTLVKFLPLMIGFALSVPSGL
SIWLNWNLSTAAQOYLKRLKGAKRNMDENASKIISAGRAKRSIAQPDADGERFOL
KDOEKRSKKNKAVAKDTVELVEESQSESGSDSEEARREGALASSTSKLPEVGO
RRSKRSRRKRY"
complement(30388..30418)
/repeat_region="AT_rich"
/tpc_family="AT_rich"
/rpc_family="AT_rich"

repeat_region complement(30388..30418)
repeat_region 31682..31794
/rpc_family="AT_rich"

Alignment Scores:
Pred. No.: 0.629 Length: 107097
Score: 13.00 Matches: 13
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.18% Indels: 0
DB: 8 Gaps: 0

US-09-397-967-16 (1-1099) x AC005727 (1-107097)
QY 1003 Setaspaalttrserpheclyalvalleutyrgluleu 1015
|||||
Db 98537 AGTGATGTATGAGCTTGGGCTAGTGTATGAACTG 98499
|||||

RESULT 66
AC005315/c 108393 bp DNA linear PLN 11-MAR-2002*
LOCUS Arabidopsis thaliana chromosome 2 clone T914 map m154, complete
DEFINITION sequence.
ACCESSION AC005315
VERSION AC005315.3 GI:20197356
KEYWORDS HTG
SOURCE Arabidopsis thaliana.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 108393)
Shen,M., Ronning,C.M., Fraser,C.M., Somerville,C.R. and Venter,J.C.
REFERENCE Rounsley,S.D., Lin,X., Kaul,S., Shea,T.P., Fujii,C.Y., Mason,T.M.,
AUTHORS Journal Title
JOURNAL Submitted (09-MAR-2000) The Institute for Genomic Research, 9712
REFERENCE Medical Center Dr., Rockville, MD 20850, USA
AUTHORS Town,C.D. and Kaul,S.

```

```

TITLE Direct Submission
JOURNAL Submitted (27-FEB-2002) The Institute for Genomic Research, 9712
COMMENT Medical Center Dr., Rockville, MD 20850, USA, cdromeltgr.org
FEATURES On Apr 18, 2002 this sequence version replaced gi:598462.
SOURCE Location/Qualifiers
1..108393
/organism="Arabidopsis thaliana"
/cultivar="Columbia"
/db_xref="taxon:3702"
/chromosome="2"
/map="m154"
/clone="T914"
1..15649
/feature="overlap with BAC clone F8N16
(AAC005727:91449..107097)."
1108..1131
/rpc_family="TAAA)n"
complement(3674..3724)
/rpc_family="AT_rich"
4287..6301
/gene="At2g28930"
/feature="T914.1: contains a protein kinase domain profile
(PDOC00100)"
join(<4287..4329,4769..5088,5179..5314,5404..5546,
5623..5746,5829..6301)
/gene="At2g28930"
join(4287..4329,4769..5088,5179..5314,5404..5546,
5623..5746,5829..6301)
/gene="At2g28930"
/codon_start=1
/product="putative protein kinase"
/protein_id="AAC33221.1"
/db_xref="GI:3461835"
/translation="MGICLSAQIKRAVSGASPKYSSSANDSLSSKSSVIRNPRT
EGEILQSPNLKSFTRAELEKATRNRPDSVLGEGFGSVFEGWIDEQTLTKSKPGTG
VTAVAKLNQDMQGEWLVLAENVYLGQSHNVLKLYCLEDEHRTSLNIEFMRGSL
ENHLERRSGSYOPSLWTLRLVVALGAAGLAFLHNAETSVYRDKTSNIYLDSEYNA
KLSPGGLAKDPTGDKSHVSTRIMGTYGAPEYLATGHTLTKSDVYSYGVALLEVIS
GRRAVDKRNRPGEOKLVEMARPLANKRLRFRVIDNRLODQYSMECKKCVATLRLC
TEFKLRPNMNEVYSHLEIOTLNAGCRNIDMQRNRKRSDSVALNOKNACFARO
TAVCIATAYRPSDSPLEFY"
complement(6389..8414)
/gene="T914.2: contains a protein kinase domain profile
(PDOC00100)"
complement(join(<6589..7152,7222..7344,7425..7688,
7977..8414))
/gene="At2g28940"
complement(join(6589..7152,7222..7344,7425..7688,
7977..8414))
/gene="At2g28940"
/codon_start=1
/product="putative protein kinase"
/protein_id="AAC33222.1"
/db_xref="GI:3461836"
/translation="MKCFHFTNGDKRRTTITIEBVGSGDSVYSRASHLSMARSLVAS
STSDPTPRSEFSDMSFSEPRLETFPKLSORMTGLVPEMDLKVFFKELKATATGF
NRGLLIGEGGCGYRGVVDSDSNGDSKTIINAVKOLNROGLOGHKEWNEVNFUCY
VNHPIYKINGCYCADDERGCMRLYELIMCNKSLDEHLVGRVSVSLPMMRKLIAQ
DAAGGLAVLEHMDQILTFDFKSSNILLDERFAKLSDFLAQCGPBEGLGVHVSIV
VGTVGAYAPAEVQGLTKARSDWSFVLYELITGRADVDRNRPREGOKLLEWKPY
VSDSKRPHLLVDPRLGEOYCKMSVQVAALANKCKOPKSRPKMSSEVYSLGRITD
EEANVPYPAVDADEETIIRKALNCESEPELKKQSSYRRKVLVDLRDMNLKSLSKID
WRNWTPELVRTW"
8721..8765
/rpc_family="AT_rich"
10356..10398
/rpc_family="AT_rich"
complement(11426..13331)
/gene="At2g28950"
/feature="T914.3: identical to GB U30480: supported by full
length cDNA: Ceres: 17914"
complement(join(11426..11885,12560..12878,13074..>13331))

```

```
repeat_region
CDS
/gene="At2g28950"
/complement(11460..11490)
/rpt_family="AT_rich"
complement(join(11576..11885,12560..12878,13074..13218))
/gene="At2g28950"
/codon_start=1
/product="expansin AtEx6"
/protein_id="AAC33223.1"
/db_xref="GI:3461837"
/translocation="MAMIGLIVLSTTTLLALSEARIPGVYNGGEMTAHATFYGGSDA
SGTMGACGYNLSGCGYNTALSTALFNNGSCGACCELCACDPKCHGSCSPSI
FITATNCPNPEAFNDNGGWCNPRPHDLAMPFLKIAEYRAIGIVPFRPDCPK
RGGITRCPNPEAFNIVLNTNACAGNIVRGVYGTHTSMWTGSRNMGQMOSNVLY
GQSLSFVYSSDRSSSTSMNAPNMFQGTFFMCKNRV"
repeat_region
16013..16090
/rpt_family="AT_rich"
17619..17647
/rpt_family="CAACAC)n"
complement(17650..22083)
/gene="At2g28960"
/note="T914.4"
complement(join(17650..18202,18290..18479,18556..18618,
18720..18846,18923..19155,19260..19417,19497..19574,
19678..19749,19937..20008,20105..20249,20591..21063,
21158..21686,21993..22083))
/gene="At2g28960"
complement(join(17794..18202,18290..18479,18556..18618,
18720..18846,18923..19155,19260..19417,19497..19574,
19678..19749,19937..20008,20105..20249,20591..21063,
21158..21686,21993..22083))
/gene="At2g28960"
/codon_start=1
/product="putative receptor-like protein kinase"
/protein_id="AAC33224.1"
/db_xref="GI:3461838"
/translocation="WEGRRQRLVIFRGNALITTHVQAQPPDQKGFSLDQGLVNES
PYDPRTGLTFSSADPFLSLGRFAGDDNVIYKQYDLDFPFGJNCLNLYEKG
INYLIRAGFGYNGDLNVPKFLDHPVNMWIVADLEGDEITLITTYTLLQICL
VKGTSTIPMISTLELRANDSYLDFGLIYRAVSSSTGTRPDDIPRKMD
HNEFETVNTLNVRSSPQVPEAVSRMGITPENALPREFVSLDDSDGKVVAV
HFAEQALRGNETREDIEDIIOASVSPMIDKAPYLNIPKSGSGCYLKLVR
TPRSTLPLPLSIAIEKRVYDPPYAPETNNDVAAKMDIAPGKIMISMGQDPCVPEL
KWEIDKCYSTNKSTPRLITSLDLSRDKGYAPFOUTELRKIDLSNNSFTGQVE
FLASKRSIINLNMNDLTGLPKLLDREKNGKLTQGNKRLCNDSSCKNNQOTY
IVPVYASVAVLIILAVLILVLEKRRPQVDSLPYVHGHPNPSIFGTQKRTYS
EVEALTDNFERVILGGGVYHGIINTOPLAYLISOSSVOGKEKKAVERLLRYS
HHVNVSLVGYCODEESNALLXEYAPNGDLAKDQLSERSGSPKMSRLKIVETAGQI
EYHTGCKPWHHRDYKTKTNILLDEHPQAKLADQLSERSGSPKMSRLKIVETAGQI
LDPEYTRNRLNEKSDVYSGVYLEITTSRPIVQOTREKPIAMWGYMLTKGDIEN
VDPRLNDYEFTSVKALDIAMSCVNSSEKRPMSOVNTLKKOCLILENSKRYRE
DMGSRSSVEMSTSFTEINPKAR"
repeat_region
22868..22909
/rpt_family="AT_rich"
22951..22971
/rpt_family="AT_rich"
23635..23655
/rpt_family="AT_rich"
23635..23655
/gene="At2g28970"
/note="T914.5"
join(23655..23724,23844..24035,24325..24379,24788..25263,
25348..25492,25589..25660,25758..25829,25929..26006,
26187..26344,26421..26671,26769..26895,26994..27059,
27203..27382,27491..27899)
/gene="At2g28970"
join(23655..23724,23844..24035,24325..24379,24788..25263,
25348..25492,25589..25660,25758..25829,25929..26006,
26187..26344,26421..26671,26769..26895,26994..27059,
27203..27382,27491..27899)
/gene="At2g28970"
/codon_start=1
/product="putative receptor-like protein kinase"
/protein_id="AAC33225.1"
/db_xref="GI:3461839"
/translocation="MMSHLLAIIIGFAVIVGAKOKEGFIISDCGPPIESPYDPST
```

```
GLTFSTDSFTEIOTGESGVKDELNKKIFRRPYLTLRYFPEGRKNCISLRNSFRVHCSTSD
SEIRYDDSDRWVWYPPFSSSEYITTSININNSDPEELIPKALKAATPKNAAVLI
ITKPPRANAERYFYIAPRAIOTLANREPRDVKPKNFSAFSPTKLELTFSTS
GPVQCDSDGNCNLDVTPNSTLPLNLNALEATIIFFPOLESTLSVNAIKINKAYR
LSTKTSWGDPCDLPQELSWENLRCSYSTNSSTPPKIIISLNASAGLSTGDSVQNLTOI
QELDLSNNSLTLGLVPSFLANIKSLSLDLSGNNTSVQTLDRKEGLVLRKEBN
ELCKFSKSCNPKRRKGLLVPIYASISSVLIVIVVAFVYLRKKMKSQAOAPSLEVE
DVGAKHSESSSEVSKKIRFAYFEVQEMTNMFQVLEGGEGVYVHAGCNGTOOAVAKL
LSOSSQSGYKHFKAEEVLLMRVHKMLVSLVGYCDGSDHALIYEYMPNGDLKOHLSG
KRGCFVLSHESRIRAVADALGLEVYHTGCKPPMVRDITKSTIILLDERFOAKLADFG
LSSSFPLENTEHSTVYAGTPTGLDEPYIOTNNLTERKSDVYSGVYLEITTRNPTIO
OSREKPHLEWVGFIVRTGDTGINVDPNLHGAYDVGSVMKAITELANSCVNAISSRRPS
MSQVSLKXCKVISENSRTGSESNMSSIEFSMGIDVEVLPKAR"
repeat_region
28787..28843
/rpt_family="AT_rich"
28870..28926
/rpt_family="AT_rich"
28870..28926
/rpt_family="AT_rich"
complement(28945..28964)
/rpt_family="(A)n"
complement(29072..34092)
/gene="At2g28980"
/note="T914.6"
complement(join(29072..32065,32249..32562,32811..34092))
/gene="At2g28980"
complement(join(29072..32065,32249..32562,32811..34092))
/gene="At2g28980"
/codon_start=1
/product="putative non-LTR retroelement reverse
transcriptase"
/protein_id="AAC33226.1"
Alignment Scores:
Pred. No.: 0.635 Length: 108393
Score: 13.00 Matches: 13
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.18% Indels: 0
Gaps: 0
US-09-397-967-16 (1-1099) x AC005315 (1-108393)
QY 1003 Seraspvalatrpserphoglyvalyvalleutyrgluleu 1015
|||||
Db 7089 AGTGATGATGAGAGCTTGAGGCTAGTGTTGATGACTG 7051
RESULT 67
LOCUS I44512 151 bp DNA linear PAT 07-OCT-1997
DEFINITION Sequence 11 from patent US 5635177.
ACCESSION I44512.1 GI:2469225
VERSION I44512.1
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 151)
AUTHORS Bennett,B.D., Goeddel,D. and Matthews,W.
TITLE Protein tyrosine kinase agonist antibodies
JOURNAL Patent: US 5635177-A 11 03-JUN-1997;
FEATURES
source
1..151
/organism="unknown"
BASE COUNT 37 a 44 c 41 g 29 t
ORIGIN
Alignment Scores:
Pred. No.: 0.0311 Length: 151
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
```

DB: 6 Gaps: 0

US-09-397-967-16 (1-1099) x I44512 (1-151)

OY 941 VALHISARGASPLEUALAALARGASNIILEUVAL 952
 |||||||
 DB 1 GTGCACAGAGATCTCGCGCTCGAACAATCTCTGTC 36

RESULT 68
 ABO12713 187 bp mRNA linear INV 10-APR-1998
 LOCUS Antihocidaris crassispina mRNA for Csk-type protein tyrosine kinase,
 DEFINITION partial cds.

ABO12713.1 GI:3043753
 VERSION Csk-type protein tyrosine kinase; ACNRTK5.
 KEYWORDS Antihocidaris crassispina cDNA to mRNA.
 SOURCE Antihocidaris crassispina
 ORGANISM Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
 Echinoidea; Euechinoidea; Echinacea; Echinoidea; Echinometridae;
 Anthocidaris.
 1 (sites)
 REFERENCE 1 (sites)
 AUTHORS Sakuma, M., Onodera, H., Suyemitsu, T. and Yamasu, K.
 TITLE The protein tyrosine kinases of the sea urchin Anthocidaris
 crassispina
 JOURNAL Zool. Sci. 14 (6), 941-946 (1997)
 MEDLINE 98181253
 REFERENCE 2 (bases 1 to 187)
 AUTHORS Yamasu, K.
 TITLE Direct Submission
 JOURNAL Submitted (30-MAR-1998) Kyo Yamasu, Saitama University, Faculty of
 Science, Department of Regulation Biology, 255 Shimo-Okubo, Urawa,
 Saitama 338-8570, Japan (E-mail:kyamasu@satl.saitama-u.ac.jp,
 Tel:+81-48-858-3417, Fax:+81-48-858-3698)
 FEATURES
 source
 1..187
 /organism="Antihocidaris crassispina"
 /db_xref="taxon:7629"
 1..187
 /gene="ACNRTK5"
 <1..>187
 /gene="ACNRTK5"
 /codon_start=1
 /product="Csk-type protein tyrosine kinase"
 /protein_id="BA025541.1"
 /db_xref="GI:3043754"
 /translation="VHRDLARNILVSGDIKAVSDGLSGEATLNOEGKFKPIKMTA
 PEALRRKEFSMSDVWST"
 CDS
 gene
 48 a 44 c 49 g 46 t

BASE COUNT 48 a 44 c 49 g 46 t

ORIGIN

Alignment Scores:
 Pred. No.: 0.0369 Length: 187
 Score: 12.00 Matches: 12
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.09% Indels: 0
 DB: 3 Gaps: 0

US-09-397-967-16 (1-1099) x ABO12713 (1-187)

OY 941 VALHISARGASPLEUALAALARGASNIILEUVAL 952
 |||||||
 DB 1 GTGCACAGAGATCTCGCGCTCGAACAATCTCTGTC 36

RESULT 69
 HYDTRYKIND 201 bp mRNA linear INV 26-APR-1993
 LOCUS Hydra vulgaris tyrosine kinase mRNA, partial cds.
 DEFINITION M64614
 ACCESSION M64614.1 GI:159281
 VERSION tyrosine kinase.
 KEYWORDS Hydra vulgaris, cDNA to mRNA.
 SOURCE

ORGANISM Hydra vulgaris
 Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydrozoa; Anthomedusae;
 Hydridae; Hydra.
 1 (bases 1 to 201)
 REFERENCE Steele, R.E.
 AUTHORS Unpublished (1991)
 JOURNAL
 FEATURES
 source
 1..201
 /organism="Hydra vulgaris"
 /db_xref="taxon:6087"
 /dev_stage="adult polyp"
 <1..>201
 /note="putative"
 /codon_start=1
 /protein_id="AAA29221.1"
 /db_xref="GI:159282"
 /translation="VHRDLARNILVSGAGKVKISDFGLTRKINDELNVMSKKRRRLP
 VKWMSVEAIFDQLTFSFSDVWTF"
 CDS
 BASE COUNT 65 a 30 c 47 g 59 t

ORIGIN

Alignment Scores:
 Pred. No.: 0.0391 Length: 201
 Score: 12.00 Matches: 12
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.09% Indels: 0
 DB: 3 Gaps: 0

US-09-397-967-16 (1-1099) x HYDTRYKIND (1-201)

OY 941 VALHISARGASPLEUALAALARGASNIILEUVAL 952
 |||||||
 DB 1 GTACATCGGATCTGCTGCAAGAAACATTTTGGTC 36

RESULT 70
 H006874S11 206 bp DNA linear PRI 29-JAN-2001
 LOCUS Homo sapiens megakaryocyte-associated tyrosine kinase (MATE) gene,
 DEFINITION exon 11.
 accession . S75159.1 GI:896215
 VERSION S75159
 KEYWORDS 11 of 13
 SEGMENT
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 206)
 REFERENCE 1 (bases 1 to 206)
 AUTHORS Avraham, S., Jiang, S., Ota, S., Fu, Y., Deng, B., Dowler, L.L.,
 White, R.A. and Avraham, H.
 TITLE Structural and functional studies of the intracellular tyrosine
 kinase MATE gene and its translated product
 JOURNAL J. Biol. Chem. 270 (4), 1833-1842 (1995)
 MEDLINE 95130565
 PUBMED 7530249
 REMARK Genbank staff at the National Library of Medicine created this
 entry [NCBI g1bbsq 159439] from the original journal article.
 This sequence comes from Fig. 5.
 FEATURES
 source
 1..206
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 6..201
 /gene="MATE"
 /number=11
 exon
 BASE COUNT 43 a 64 c 72 g 27 t

ORIGIN

Alignment Scores:
 Pred. No.: 0.0399 Length: 206
 Score: 12.00 Matches: 12
 Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.09% Indels: 0
 DB: 9 Gaps: 0

US-09-397-967-16 (1-1099) x H006874511 (1-206)

QY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
 |||||||
 Db 49 GTGCACCGGACCTGGCCCGGACACATCTGTGTC 84

RESULT 71
 AX203083 214 bp DNA linear PAT 30-AUG-2001
 LOCUS Sequence 4 from Patent WO0152892.
 AX203083
 ACCESSION AX203083.1 GI:15392440
 VERSION
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 214)
 AUTORS Vassios,G.
 TITLE Jak/stat pathway inhibitors and the uses thereof
 JOURNAL Patent: WO 0152892-A 4 26-JUL-2001;
 GENZYME CORPORATION (US)
 FEATURES
 source Location/Qualifiers
 1..214
 /organism="Homo sapiens"
 /db_xref="taxon:9606"

BASE COUNT 34 a 74 c 62 g 36 t 8 others
 ORIGIN

Alignment Scores:
 Pred. No.: 0.0411 Length: 214
 Score: 12.00 Matches: 12
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.09% Indels: 0
 DB: 6 Gaps: 0

US-09-397-967-16 (1-1099) x AX203083 (1-214)

QY 938 ArgArgCysValHisArgAspLeuAlaAlaArgAsn 949
 |||||||
 Db 128 CCGCCGTCGTCGACCGCGACCTGGCCCGGAC 163

RESULT 72
 AB038987 302 bp mRNA linear VRT 28-AUG-2001
 LOCUS Gallus gallus EphA9 mRNA for protein tyrosine kinase EphA9, partial
 DEFINITION cds.
 AB038987
 ACCESSION AB038987.1 GI:15375362
 VERSION
 KEYWORDS protein tyrosine kinase EphA9.
 SOURCE Gallus gallus Stage X embryos Blastodermal Cells cDNA to mRNA.
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 Phasianinae; Gallus.
 1
 REFERENCE 1
 AUTHORS Sasaki,E., Sakurai,M. and Etches,R.J.
 TITLE Expression of Protein Tyrosine Kinase in Chicken Blastodermal Cells
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 302)
 AUTORS Sasaki,E., Sakurai,M. and Etches,R.J.
 TITLE Direct Submission
 JOURNAL Submitted (25-FEB-2000) Erika Sasaki, University of Guelph,
 Department of Food Science, Guelph, Ontario, N1G 2W1, Canada
 (E-mail:enakaminuoguelph.ca, Tel:1-519-824-4120(ex.2259),
 Fax:1-519-824-6631)
 FEATURES
 source Location/Qualifiers
 1..302

gene
 CDS
 1..302
 /gene="EphA9"
 <1..>302
 /gene="EphA9"
 /codon_start=3
 /product="protein tyrosine kinase EphA9"
 /protein_id="BAB63911.1"
 /db_xref="GI:15375363"
 /translation="NGALDTELENEERKFPVQVLYNMLQGIASGMTYLSNNYVHRL
 AARNILVTRSLQKVSDFSLRIENDAGTYETKGIPIRWTFPAIAHRIFTS"

BASE COUNT 80 a 75 c 76 g 71 t
 ORIGIN

Alignment Scores:
 Pred. No.: 0.0542 Length: 302
 Score: 12.00 Matches: 12
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.09% Indels: 0
 DB: 5 Gaps: 0

US-09-397-967-16 (1-1099) x AB038987 (1-302)

QY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
 |||||||
 Db 120 GTGCACCGGACCTGGCTGCTCGACACATCTGTGTA 155

RESULT 73
 CHKTYKDA 373 bp DNA linear VRT 28-APR-1993
 LOCUS CHKTYKDA
 DEFINITION Chicken DNA that hybridizes to v-fps tyrosine kinase domain of the
 Rous sarcoma virus.
 ACCESSION M12257
 VERSION M12257.1 GI:212845
 KEYWORDS
 SOURCE Chicken DNA, clone pTKR11.
 ORGANISM Gallus gallus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 Phasianinae; Gallus.
 REFERENCE 1 (bases 1 to 373)
 AUTORS Foster,D.A., Levy,J.B., Daley,G.O., Simon,M.C. and Hanafusa,H.
 TITLE Isolation of chicken cellular DNA sequences with homology to the
 region of viral oncogenes that encodes the tyrosine kinase domain
 JOURNAL Mol. Cell. Biol. 6 (1), 325-331 (1986)
 MEDLINE 87064304
 PUBMED 3023834
 COMMENT Draft entry and computer readable copy of sequence [1] kindly
 provided by D.A.Foster, 30-MAY-1986.
 FEATURES
 source Location/Qualifiers
 1..373
 /organism="Gallus gallus"
 /db_xref="taxon:9031"

BASE COUNT 90 a 90 c 104 g 89 t
 ORIGIN

Alignment Scores:
 Pred. No.: 0.0642 Length: 373
 Score: 12.00 Matches: 12
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.09% Indels: 0
 DB: 5 Gaps: 0

US-09-397-967-16 (1-1099) x CHKTYKDA (1-373)

QY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
 |||||||
 Db 60 GTGCACCGGACCTGGCTGCTCGGACACATCTGTGTC 95

RESULT 74
AR053289 738 bp DNA linear PAT 29-SEP-1999
LOCUS AR053289
DEFINITION Sequence 8 from patent US 5834208.
ACCESSION AR053289
VERSION AR053289.1 GI:5978151
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 738)
AUTHORS Sakano,S.
TITLE Tyrosine kinase
JOURNAL Patent: US 5834208-A 8 10-NOV-1998;
FEATURES
source Location/Qualifiers
BASE COUNT 154 a 205 c 248 g 131 t
ORIGIN
Alignment Scores:
Pred. No.: 0.111 Length: 738
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 6 Gaps: 0
US-09-397-967-16 (1-1099) x AR053289 (1-738)
QY 941 VALHSAAGSAPLEAALAAAGASNLLEUVAL 952
Db 349 GTGACCGCCAGCTGCGCCGCGACATCTGTGTC 384
RESULT 75
RNO299010 790 bp DNA linear ROD 05-MAR-2002
LOCUS Rattus norvegicus partial Ret gene for receptor tyrosine kinase,
exons 14-15.
ACCESSION AJ299010
VERSION AJ299010.1 GI:10638054
KEYWORDS receptor tyrosine kinase; ret gene; RET proto-oncogene.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1
AUTHORS Matera,I., De Miguel-Rodriguez,M., Fernandez-Santos,J.M.,
Santamaria,G., Puliti,A., Ravazzolo,R., Romeo,G.,
Galera-Davidson,H. and Ceccherini,I.
TITLE cDNA sequence and genomic structure of the rat RET proto-oncogene
JOURNAL DNA Seq. 11 (5), 405-417 (2000)
MEDLINE 21226279
PUBMED 11328649
REFERENCE 2 (bases 1 to 790)
AUTHORS Ceccherini,I.
TITLE Direct Submission
JOURNAL Submitted (21-SEP-2000) Ceccherini I., Laboratorio Genetica
Molecolare, Istituto Giannina Gaslini, Largo G. Gaslini, 5 -
Genova, 16148, ITALY
FEATURES
source Location/Qualifiers
1..790
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
1..790
/gene="Ret"
<1..283
/gene="Ret"
/number=13
284..498
/gene="Ret"
exon
Intron

/number=14
/usedin-AJ298999:Ret_CDS1
/usedin-AJ298999:Ret_CDS2
/usedin-AJ298999:Ret_mRNA1
/usedin-AJ298999:Ret_mRNA2
499..591
/gene="Ret"
/number=14
592..714
/gene="Ret"
/number=15
/usedin-AJ298999:Ret_CDS1
/usedin-AJ298999:Ret_CDS2
/usedin-AJ298999:Ret_mRNA1
/usedin-AJ298999:Ret_mRNA2
715..>790
/gene="Ret"
/number=15
Intron
BASE COUNT 178 a 203 c 230 g 177 t 2 others
ORIGIN
Alignment Scores:
Pred. No.: 0.117 Length: 790
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 10 Gaps: 0
US-09-397-967-16 (1-1099) x RNO299010 (1-790)
QY 941 VALHSAAGSAPLEAALAAAGASNLLEUVAL 952
Db 595 GTACATCGAGACTTACTGCTCCAGAACATCTTGTGTC 630
RESULT 76
DR089380 877 bp mRNA linear VRT 26-JUN-1997
LOCUS Danio rerio Eph-like kinase 2 (zek2) mRNA, partial cds.
DEFINITION U89380
ACCESSION U89380
VERSION U89380.1 GI:2198796
KEYWORDS
SOURCE Danio rerio.
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
REFERENCE 1 (bases 1 to 877)
AUTHORS Bovenkamp,D.E. and Greer,P.
TITLE Novel Eph-family receptor tyrosine kinase is widely expressed in
JOURNAL the developing zebrafish nervous system
MEDLINE Dev. Dyn. 209 (2), 166-181 (1997)
PUBMED 9186052
REFERENCE 2 (bases 1 to 877)
AUTHORS Bovenkamp,D.E. and Greer,P.
TITLE Direct Submission
JOURNAL Submitted (13-FEB-1997) Department of Biochemistry, Cancer Research
Labs, Queen's University, Rm. A309 Botterell Hall, Kingston, ON K7L
3N6, Canada
FEATURES
source Location/Qualifiers
1..877
/organism="Danio rerio"
/db_xref="taxon:7955"
<1..>877
/gene="zek2"
<1..>877
/gene="zek2"
/note="Eph-like kinase 2; zek2; Eph-like receptor tyrosine
kinase"
/codon_start=1
/product="zek2"
/protein_id="AAC60222.1"

```

/db_xref="GI:2198797"
/translation="IGIGFGEVCSGRLKMPGRKEICVAIKTLKAGYTDKORDPLISE
ASIMGODHNIIRLEGVYTKCKPVMITTEYMGNSIDFLRNDGRFTVIOYVGLIR
GIASGMKYSIDMSVHRDLAARNILVNSICVSDGMSRLEEDPDAAVYTRREITG
TYOSOGSKIPRTWTAPEALITVRKSTSDSYGIVMVEVNSYGERPYWMSNODYIK
AIEBGRKLPPEPCVPSLHQLMDCKMRBAERKPSQIYNMIDKLIRNNSLKRRFG
EIARNPTLLEPSSPE"
BASE COUNT      245 a      204 c      236 g      192 t
ORIGIN

Alignment Scores:
Pred. No.:      0.127      Length:      877
Score:          12.00      Matches:     12
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    1.09%      Indels:      0
DB:              5      Gaps:      0

US-09-397-967-16 (1-1099) x DR089380 (1-877)

Oy  941 ValHisArgAspLeuAlaAlaArgAsnIleuVal 952
DB  349 GTCCATCGTGACTGCGCGCTGCAACATCTGTGTC 384

RESULT 77
MMSEK4 LOCUS      MMSEK4      1128 bp      mRNA      linear      ROD 27-APR-1995
DEFINITION M.musculus Sek-4 mRNA.
ACCESSION X76012
VERSION X76012.1 GI:460057
KEYWORDS   Sek gene.
SOURCE     Mus musculus.
ORGANISM   Mus musculus.
REFERENCE 1 Mus musculus.
AUTHORS    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
          1 (bases 1 to 1128)
          Becker N., Seltanidou, T., Murphy, P., Mattei, M.G., Topilko, P.,
          Nieto, M.A., Wilkinson, D.G., Charnay, P. and Gilardi-Hebenstreit, P.
          Several receptor tyrosine kinase genes of the Eph family are
          segmentally expressed in the developing hindbrain
          Mech. Dev. 47 (1), 3-17 (1994)
          2 (bases 1 to 1128)
          Becker, N.
          Direct Submission
          Submitted (11-OCT-1993) N. Becker, Unite Inserm U368 Ecole Normale,
          Superieure, Laboratoire de Biologie Moleculaire du Developpement-
          Patrick Charnay, U368 INSERM 46 Rue d'Ulm 75005 Paris, FRANCE
          Location/Qualifiers
          1..1128
             /organism="Mus musculus"
             /strain="BALB/c"
             /db_xref="taxon:10090"
             /chromosome="16"
             /map="B1-B4"
             /clone="20"
             /tissue_type="brain"
             /gene="Sek-4"
             /gene="Sek-4"
             /c1..828
             /gene="Sek-4"
             /codon_start=1
             /protein_id="CAA53599.1"
             /db_xref="GI:460058"
             /db_xref="MIM:104770"
             /db_xref="SPTREMBL:O62214"
             /translation="QLNDGFTYIQVGMIRGIAAGMKYISEMNVYHRDLAARNILVN
             SNLVCKVSDGSLRFLDEDPDPTVSSLGKIPRTWTAPEALITVRKSTSDSYGIVMVEVNSYGERPYWMSNODYIK
             IVMVEVNSYGERPYWMSNODYIKAVEDYRLPPPCPTALHQLMDCKMRBAERKPSQIYNMIDKLIRNNSLKRRFG
             KFSQIVNTDLKILRNASLKVITASAPSGNSQPLDRTVDTFTTGVGMIDAIKNGR
             YKRSFVGAGFASPDIVAQMATAEDLRLRGVTLAGHQKILCSIDMRLOMNTLPVGV"
             /misc_feature
             1..828

```

```

/gene="Sek-4"
/note="Intra cellular region"
BASE COUNT      245 a      326 c      308 g      249 t
ORIGIN

Alignment Scores:
Pred. No.:      0.155      Length:     1128
Score:          12.00      Matches:     12
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    1.09%      Indels:      0
DB:              10      Gaps:      0

US-09-397-967-16 (1-1099) x MMSEK4 (1-1128)

Oy  941 ValHisArgAspLeuAlaAlaArgAsnIleuVal 952
DB  94  GTGCACCGTGACCTTGCTGCCGCAACATCTGTGTC 129

RESULT 78
LOCUS      AB025542      1153 bp      mRNA      linear      VRT 18-NOV-1999
DEFINITION Lampetra reissneri mRNA for Ephb, partial cds.
ACCESSION AB025542
VERSION AB025542.1 GI:6002426
KEYWORDS   Ephb.
SOURCE     Lampetra reissneri larva cDNA to mRNA.
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
          Petromyzontiformes; Petromyzontidae; Lethenteron.
          1 (sites)
          Suga, H., Hoshiyama, D., Kuraku, S., Katoh, K., Kubokawa, K. and
          Miyata, T.
          Protein tyrosine kinase CDNAs from amphioxus, hagfish, and lamprey:
          isoform duplications around the divergence of cyclostomes and
          gnathostomes
          J. Mol. Evol. 49 (5), 601-608 (1999)
          2 (bases 1 to 1153)
          Suga, H.
          Direct Submission
          Submitted (31-MAR-1999) Hiroshi Suga, Faculty of Science, Kyoto
          University, Department of Biophysics, Kitashirakawa-oiwa-cho,
          Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
          (E-mail:sugahbiophys.kyoto-u.ac.jp, Tel:+81-75-753-4224,
          Fax:+81-75-753-4223)
          Location/Qualifiers
          1..1153
             /organism="Lethenteron reissneri"
             /db_xref="taxon:7753"
             /dev_stage="larva"
             /c1..1063
             /codon_start=2
             /product="Ephb"
             /protein_id="BA084732.1"
             /db_xref="GI:6002427"
             /translation="EVCSGRLLKPKREMAVAIKTLKAGYTERORDELSEASIMGOF
             DHPVYHIREGVVTSRPVMTTEFMENSLDSFRONDGFTVIOYVGMIRGIAAGMK
             YLADMSYVHRDLAARNILVNSICVSDGMSRLEEDPDPTVSSAMGKIPRTW
             APEALITVRKSTSDSYGIVMVEVNSYGERPYWMSNODYIKAVEDYRLPPPCPTALHQLMDCKMRBAERKPSQIYNMIDKLIRNNSLKRRFG
             YKRSFVGAGFASPDIVAQMATAEDLRLRGVTLAGHQKILCSIDMRLOMNTLPVGV"
             /misc_feature
             1..828

```


US-09-397-967-16 (1-1099) x AB025542 (1-1153)

Oy 941 VALHISARGASPLEUALAALARGASNIJLEUVAL 952
 Db 329 GTGCATCGAGATCTGCGACGACGACATCTCTGTC 364

RESULT 79

LOCUS HYDTRYRKINA 1186 bp mRNA linear INV 08-JUL-1999
 DEFINITION Hydra vulgaris putative tyrosine kinase mRNA, complete cds.
 ACCESSION M64611
 VERSION M64611.1 GI:159275
 KEYWORDS
 SOURCE Hydra vulgaris.
 ORGANISM Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydrozoa; Anthomedusae;
 Hydridae; Hydra.

REFERENCE 1 (bases 1 to 1186)
 AUTHORS Krolher, M., Reidling, J.C. and Steele, R.E.
 TITLE A gene whose major transcript encodes only the substrate binding domain of a protein-tyrosine kinase

JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1186)
 AUTHORS Steele, R.E.

JOURNAL Unpublished
 REFERENCE 3 (bases 1 to 1186)
 AUTHORS Krolher, M., Reidling, J.C. and Steele, R.E.
 TITLE Direct Submission
 JOURNAL Submitted (01-MAY-1991) Department of Biological Chemistry, University of California, 240D Medical Sciences I, Irvine, CA 92697-1700, USA

FEATURES
 source Location/Qualifiers
 1..1186

/organism="Hydra vulgaris"
 /db_xref="taxon:6087"
 /dev_stage="adult polyp"
 110..823
 /codon_start=1
 /product="putative tyrosine kinase"
 /protein_id="AA029219.1"
 /db_xref="GI:159276"

/translation="MIGCSTLQKPICLVEMENDLQFLKNSRTKLITSLYKSS
 VIFNPNOMPEEYSSIKLVHRLAARNILVAGADIKYISDFGLTRVNDIYMG
 NNRLPIKIMSEATFEDRTFTSYDVMAVGIYLPETVILGSPYSISNRELPLKS
 CYRMERPNCSMPCDMLNCHNEPDLQRPFTTKLRETFEELMSGNTLNEISEEC
 TTYNPPSSHSNNDLVV"

BASE COUNT 411 a 154 c 196 g 425 t
 ORIGIN

Alignment Scores:

Pred. No.: 0.162 Length: 1186
 Score: 12.00 Matches: 12
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.09% Indels: 0
 DB: 3 Gaps: 0

US-09-397-967-16 (1-1099) x HYDTRYRKINA (1-1186)

Oy 941 VALHISARGASPLEUALAALARGASNIJLEUVAL 952
 Db 302 GTACATCGAGATCTGCTCTAGAAATATTATTAGTA 337

RESULT 80

LOCUS HUMERK 1225 bp mRNA linear PRI 23-JUN-1999
 DEFINITION Human mRNA for large erk kinase.
 ACCESSION D14717
 VERSION D14717.1 GI:285916
 KEYWORDS large erk kinase.
 SOURCE Homo sapiens gastric carcinoma, CDNA to mRNA, clone H1.
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (sites)
 AUTHORS Chan, J. and Watt, V.M.
 TITLE eek and erk, new members of the eph subclass of receptor protein-tyrosine kinases
 JOURNAL Oncogene 6 (6), 1057-1061 (1991)
 MEDLINE 91296384

REFERENCE 2 (sites)
 AUTHORS Iwase, T., Tanaka, M., Suzuki, M., Naito, Y., Sugimura, H. and Kino, I.
 TITLE Identification of protein-tyrosine kinase genes preferentially expressed in embryo stomach and gastric cancer
 JOURNAL Biochem. Biophys. Res. Commun. 194 (2), 698-705 (1993)
 MEDLINE 9343925
 REFERENCE 3 (bases 1 to 1225)
 AUTHORS Iwase, T.

JOURNAL Unpublished
 COMMENT Submitted (23-MAR-1993) to DDBJ by:

Toshio Iwase
 First Department of Pathology
 Hamamatsu University
 School of Medicine
 3600 Handa-cho, Hamamatsu
 Shizuoka 431-31
 Japan

Phone: 053-435-2220
 Email: toiwase@dbj.nig.ac.jp
 Fax: 053-435-2225

FEATURES

source

Location/Qualifiers
 1..1225

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /tissue_type="gastric carcinoma"
 1..1225
 /gene="large erk"
 <1..1044
 /gene="large erk"
 /function="protein-tyrosine kinase"
 /codon_start=1
 /product="large erk kinase"
 /protein_id="BA03537.1"
 /db_xref="GI:285917"

/translation="LKDPKKEITVAIKTKLGSYTEKQRDELSEKSGQFDHPNVI
 HLESVKSTVMTITTEPMENGLDSELRDQGFYIQLVGLRGIAAGKRYLADNN
 YRHTDLAARNILVSNLVCKVSDGLSFLEDDYSDPTYSALGCKIPIMTAPALQ
 YRHTSDASDVSYGIVMEVSYGERPMDTNDGVINALIODYRLPPMPCPALHO
 LMLDCMOKDRNHRPKFGQIVNTLDMKTRNPSLKAMPLSSGINTPLDRTIPYTSF
 NTVDEMLKAIKMGQYKESFANAGFTSDVYSQMMEDILRVGTVLAGHOKKILNSIOV
 MRAQMNOISVEV"

unsure

/gene="large erk"
 /citation=[1]
 /replac="gc"
 220..225
 /gene="large erk"
 /citation=[1]
 /replac="gtagg"

BASE COUNT 283 a 386 c 329 g 227 t
 ORIGIN

Alignment Scores:

Pred. No.: 0.166 Length: 1225
 Score: 12.00 Matches: 12
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.09% Indels: 0
 DB: 9 Gaps: 0

US-09-397-967-16 (1-1099) x HUMERK (1-1225)

Oy 941 VALHISARGASPLEUALAALARGASNIJLEUVAL 952
 Db 310 GTTCACGCTGAGCTGCGCCGACACATCTCTGTC 345

```

RESULT 81
MMSEK2
LOCUS      M.musculus Sek-2 mRNA.                1294 bp      mRNA      linear      ROD 10-MAR-2001
DEFINITION X76010.X57343
ACCESSION  X76010.1 GI:460053
KEYWORDS   Sek gene.
SOURCE     house mouse.
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE  1 (bases 1 to 1294)
AUTHORS   Becker, N., Seitanidou, T., Murphy, P., Mattei, M.G., Topilko, P.,
            Nieto, M.A., Wilkinson, D.G., Charnay, P. and Giliardi-Hedenstreit, P.
            Several receptor tyrosine kinase genes of the Eph family are
            segmentally expressed in the developing hindbrain
            Mech. Dev. 47 (1), 3-17 (1994)

JOURNAL    95034306
MEDLINE    7947319
PUBMED     2 (bases 1 to 1294)
REFERENCE  Becker, N.
            Direct Submission
            Submitted (11-OCT-1993) N. Becker, Unite Inserm U368 Ecole Normale,
            Supérieure, Laboratoire de Biologie Moléculaire, du Développement-
            Patrick Charnay, U368 INSERM 46 Rue d'Ulm 75005 Paris, FRANCE

FEATURES   location/Qualifiers
SOURCE     1..1294
            /organism="Mus musculus"
            /strain="C57 BL"
            /db_xref="taxon:10090"
            /chromosome="4"
            /map="D.E"
            /clone="1"
            /tissue-type="total embryo"
            /clone_lib="Brightle HOGAN"
            /dev_stage="embryo, 8.5 days"
            1..1281
            /gene="Sek-2"
            <1..1281
            /gene="Sek-2"
            /codon_start=1
            /protein_id="CA053597.1"
            /db_xref="GI:460054"
            /db_xref="SWISS-PROT:Q03145"
            /translation="LAGVLFIRRRRLRAROSSEDEVRFKSKDLKPLKTYVDPHTY
            EDPNQAVIKFTTEIHPSCVARKQVIGAGEFEVYKGTAKSSGKEIPVALIKTKAGY
            TEKORVDFLSASIMGSPSHNIIRLEGVSKYKPMITTEMGALDKFLREKDG
            FSVLDVGLKGLASGKTYLNNMTVHDLARNILVNSLVKVSDEGLSRVLEDDP
            EAYVTSGKLPITRWTAPEALSYRKFTSASVWSIGIVMEVMTIGEPYELSNHEY
            MKAINDFRLPTMDCPSAIVQLMMQCQOERSRPRFADIVSLDKLIRAPDSIKTL
            ADPFRVIRLPTSGSEGVPEFTVSEMLSEIKMQOYEHFMVAGYTAIEVYVMSNE
            DIKRGVRLPGHKRIAYSLGLKQVNTVIGPI"
            1..27
            /gene="Sek-2"
            /note="Trans membrane region"
            28..1281
            /gene="Sek-2"
            /note="Intra cellular region"

BASE COUNT 314 a      363 c      365 g      252 t

ORIGIN
Alignment Scores:
Pred. No.:      0.174      Length:      1294
Score:          12.00      Matches:      12
Percent Similarity: 100.00%  Mismatches:  0
Best Local Similarity: 100.00%  Indels:      0
Query Match:    1.09%      Gaps:        0
DB:            10

US-09-397-967-16 (1-1099) x MMSEK2 (1-1294)
Oy      941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952

```

```

Db      556 GTGCACCGGAGACTGGCCGCCGACACATCCTGTC 591
|||||
RESULT 82
LOCUS      AR053290                                1398 bp      DNA      linear      PAT 29-SEP-1999
DEFINITION Sequence 9 from patent US 5834208.
ACCESSION  AR053290
VERSION    AR053290.1 GI:5978152
KEYWORDS   Unknown.
SOURCE     Unknown.
ORGANISM   Unclassified.
            1 (bases 1 to 1398)

REFERENCE  Sakano, S.
            Tyrosine kinase
            Patent: US 5834208-A 9 10-NOV-1998;
            Location/Qualifiers
            1..1398
            /organism="unknown"

BASE COUNT 282 a      426 c      468 g      222 t

ORIGIN
Alignment Scores:
Pred. No.:      0.185      Length:      1398
Score:          12.00      Matches:      12
Percent Similarity: 100.00%  Mismatches:  0
Best Local Similarity: 100.00%  Indels:      0
Query Match:    1.09%      Gaps:        0
DB:            6

US-09-397-967-16 (1-1099) x AR053290 (1-1398)
Oy      941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
|||||
Db      922 GTGCACCGGAGACTGGCCGCCGACACATCCTGTC 957
|||||
RESULT 83
LOCUS      DRAJ5028                                1473 bp      mRNA      linear      VRT 27-MAY-1998
DEFINITION Danio rerio mRNA for Eph-like receptor tyrosine kinase rtk7,
            partial.
ACCESSION  AJ005028
VERSION    AJ005028.1 GI:3005904
KEYWORDS   Eph-like receptor tyrosine kinase.
            zebrafish.
SOURCE     Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
            1 (bases 1 to 1473)
            Cooke, J.E.
            Direct Submission
            Submitted (27-MAR-1998) Cooke J.E., University College London,
            Department of Anatomy and Developmental Biology, Gower Street,
            London, WC1E 6BT, UK
            2 (bases 1 to 1473)
            Cooke, J.E., Xu, Q., Wilson, S.W. and Holder N.
            Characterisation of five novel zebrafish Eph-related receptor
            tyrosine kinases suggests roles in patterning the neural plate
            Dev. Genes Evol. 206, 515-531 (1997)
            Location/Qualifiers
            1..1473
            /organism="Danio rerio"
            /db_xref="taxon:7955"
            <1..1473
            /codon_start=1
            /product="Eph-like receptor tyrosine kinase rtk7"
            /protein_id="CA06301.1"
            /db_xref="GI:3005905"
            /db_xref="SPTREMBL:O73877"
            /translation="KNSMTVDGLKPGTYYIFRYRARTDGGYNGYKGEILETSHEDM
            LAVGPNOQTIIATISVAGAVLVLVACFIVSGRRCGYIKAKODPEERKMOFOGHRY"

BASE COUNT 314 a      363 c      365 g      252 t

ORIGIN
Alignment Scores:
Pred. No.:      0.174      Length:      1294
Score:          12.00      Matches:      12
Percent Similarity: 100.00%  Mismatches:  0
Best Local Similarity: 100.00%  Indels:      0
Query Match:    1.09%      Gaps:        0
DB:            10

US-09-397-967-16 (1-1099) x MMSEK2 (1-1294)
Oy      941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952

```

KLPTRTYIHPTIEDPNQAVDFAKELIEVSNIRIERVAGIEGEGVCSGRLLRLEPSK
 EIOVAIKSLKAGYSEHORRDLSEASIMGQFDHPNIIIEGVTRCKPMIVTEYMN
 GSDPTFKHKGQFVLIQGLMGLOYSLEPMNVHDLARNILVNGNVCYK
 SDPGISRLIEDDPEAAVTRGGKIPIRMTAPEATYTRKFTSASDYSYGYMVEVYSY
 GERPYWMSNDVYKAVDEGRRLRPMPCPYVYLHLMDCNEKNSDPRKQIYNTL
 DRLIRNPSLKLQANSVAVDEDPVPEAAVNTVEMLDLIRKGOYKREHSSAGVYILDS
 VLYVSSSLDMLGVELAGHOKRKLISLQICLAAHGTQYOV

BASE COUNT 387 a 341 c 431 g 314 t
 ORIGIN

Alignment Scores:

Pred. No.: 0.192 Length: 1473
 Score: 12.00 Matches: 12
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.09% Indels: 0
 DB: 5 Gaps: 0

US-09-397-967-16 (1-1099) x DRAJ5028 (1-1473)

QY 941 VALHISARGASPLEUALAALARGASNIIEUVAL 952
 DB 769 GTGCATCGGAGATTGTGCACGACGACATTCGTGTA 804

RESULT 84

RNELK 1509 bp mRNA linear ROD 14-JUL-1995
 LOCUS Rcel mRNA for elk protein.
 DEFINITION X13411
 X13411.1 GI:56094
 elk gene; protein-tyrosine kinase.
 KEYWORDS Rattus rattus
 SOURCE Rattus rattus
 ORGANISM Rattus rattus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

REFERENCE 1 (bases 1 to 1509)

AUTHORS Letwin,K.
 TITLE Direct Submission
 JOURNAL Submitted (03-NOV-1988) Letwin K., Mount Sinai Hospital Research
 Institute, 600 University Avenue, Toronto, Ontario, M5G 1X5
 REFERENCE 2 (bases 1 to 1509)
 AUTHORS Letwin,K., Yee,S.P. and Pawson,T.
 TITLE Novel protein-tyrosine kinase cDNAs related to fps/fee and eph
 cloned using anti-phosphotyrosine antibody
 JOURNAL Oncogene 3 (6), 621-627 (1988)
 MEDLINE 94167102
 PUBMED 2485255
 COMMENT Data kindly reviewed (03-Feb-1989) by Letwin K.
 FEATURES
 SOURCE location/Qualifiers
 1..1509
 /organism="Rattus rattus"
 /strain="Wistar"
 /db_xref="taxon:10117"
 /clone="lambda B1"
 /tissue_type="brain"
 /dev_stage="adult"
 <1..1143
 /codon_start=1
 /product="elk protein"
 /protein_id="CAA31777.1"
 /db_xref="GI:56095"
 /db_xref="SWISS-PROT:P09759"
 /translation="EAVREFAKEIDVSPFKIEEVIGAGEGEVYKGRLLPGKREIY
 AITKLKAGYSEHORRDLSEASIMGQFDHPNIIIEGVTRCKPMIVTEYMN
 SELRNQDQFVLIQGLMGLOYSLEPMNVHDLARNILVNGNVCYKSDPG
 LSRYLQDTSPTTSSLGKIPYRWTPKPAIYRKTFSADVNSYGYMVEVNSFCG
 RPYWMSNDVYKAVDEGRRLRPMPCPYVYLHLMDCNEKNSDPRKQIYNTL
 MIRNPASLKYATITVAVPSQPLDRLSDPTFTVDMLSAIKWQYRDSFLTAGFT
 SIQLVTQWTSDELIRIGVTLAGHOKRKLISLSHRSVQNSPSVMA"
 23..274
 /note="catalytic domain"

misc_feature 393 a 373 c 430 g 313 t
 BASE COUNT

ORIGIN

Alignment Scores:
 Pred. No.: 0.196 Length: 1509
 Score: 12.00 Matches: 12
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.09% Indels: 0
 DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x RNELK (1-1509)

QY 941 VALHISARGASPLEUALAALARGASNIIEUVAL 952
 DB 409 GTGCACCGGACCTGGCTCTAGAACATTCGTGTA 444

RESULT 85

AR053291 1521 bp DNA linear PAT 29-SEP-1999
 LOCUS AR053291
 DEFINITION Sequence 10 from patent US 5834208.
 ACCESSION AR053291
 VERSION AR053291.1 GI:5978153
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 1521)
 AUTHORS Sakano,S.
 TITLE Tyrosine kinase
 JOURNAL Patent: US 5834208-A 10 10-NOV-1998;
 FEATURES location/Qualifiers
 source 1..1521
 /organism="unknown"

BASE COUNT

295 a 470 c 507 g 249 t
 ORIGIN

Alignment Scores:

Pred. No.: 0.197 Length: 1521
 Score: 12.00 Matches: 12
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.09% Indels: 0
 DB: 6 Gaps: 0

US-09-397-967-16 (1-1099) x AR053291 (1-1521)

QY 941 VALHISARGASPLEUALAALARGASNIIEUVAL 952
 DB 1045 GTGCACCGGACCTGGCTCTAGAACATTCGTGTA 1080

RESULT 86

XELPHRKT 1531 bp mRNA linear VRT 02-APR-1996
 LOCUS XELPHRKT
 DEFINITION Xenopus laevis eph receptor tyrosine kinase subfamily mRNA, 3' end
 of cds.
 ACCESSION L43622
 L43622.1 GI:974707
 VERSION eph-class receptor tyrosine kinase; receptor tyrosine kinase.
 KEYWORDS Xenopus laevis (clone: PL7a) gastrula cDNA to mRNA.
 SOURCE Xenopus laevis
 ORGANISM Xenopus laevis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae;
 Xenopodinae; Xenopus.
 REFERENCE 1 (bases 1 to 1531)
 AUTHORS Scales,J.B., Wining,R.S., Renaud,C.S., Shea,L.J. and Sargent,T.D.
 TITLE Novel members of the eph receptor tyrosine kinase subfamily
 expressed during Xenopus development
 JOURNAL Oncogene 11 (9), 1745-1752 (1995)
 MEDLINE 96068901
 PUBMED 7478602

location/Qualifiers
 1..1531
 /organism="Xenopus laevis"

```

/ab_xref="taxon:8355"
/clone="pl7a"
/dev_stage="gastrula"
<1..972
/codon_start=1
/product="Eph receptor tyrosine kinase"
/db_xref="GI:974708"
/translation="EFISFASIMGCFNHNPIHLEGVITNNGCVMTITEYMENGALDS
FLKNDQGTPIOLVGLMKGIAAGKILSEMSVTHDLARNITLVNSILVCKVSDGEL
SRFLQGSTDPYITSLGKIPIKMTAPEIAIKFTSADVGWIGIYVMEVVSFGER
PYWMSNOVDYINAEQDYRLPAPDCPTALHQLMDCCOARDLRFRFADIVSLDKL
IRNPASLKITTEFQISLPILDORTPHYSSFSVSSEMLHAIKMGREYEDGFRNAGPTT
FSRVNISTEDILRMGVTLAGHOKILISSIQILPSEKSTPIPDHCY"

BASE COUNT      377 a      410 c      383 g      361 t

Alignment Scores:
Pred. No.:      0.199      Length:      1531
Score:          12.00      Matches:      12
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    1.09%      Indels:      0
DB:             5      Gaps:      0

US-09-397-967-16 (1-1099) x XELEPHRTK (1-1531)

QY      941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
      |||||||
Db      232 GTGCACCGGAGATCTGCCGCGAGAAATATCTAGTG 267

RESULT 87
LOCUS      AB025541      1563 bp      mRNA      linear      VRT 18-NOV-1999
DEFINITION Eptaretus burgeri mRNA for EphA, partial cds.
ACCESSION      AB025541
VERSION      AB025541.1      GI:6002424
KEYWORDS      EphA.
SOURCE      Eptaretus burgeri adult female cDNA to mRNA.
ORGANISM      Eptaretidae; Chordata; Craniata; Hyperotretli; Myxiniiformes;
      Myxiniidae; Eptaretinae; Eptaretus.
REFERENCE      1 (sites)
      Suga,H., Hoshiyama,D., Kuraku,S., Katch,K., Kubokawa,K. and
      Miyata,T.
      Protein tyrosine kinase cDNAs from amphioxus, hagfish, and lamprey:
      isoform duplications around the divergence of cyclostomes and
      gnathostomes
      J. Mol. Evol. 49 (5), 601-608 (1999)
      20020330
      2 (bases 1 to 1563)
      Suga,H.
      Direct Submission
      Submitted (31-MAR-1999) Hiroshi Suga, Faculty of Science, Kyoto
      University, Department of Biophysics, Kitashirakawa-oiwakecho,
      Sakyo-ku, Kyoto 606-8502, Japan
      (E-mail:suga@biophys.kyoto-u.ac.jp, Tel:+81-75-753-4224,
      Fax:+81-75-753-4223)
      location/Qualifiers
      1.1563
      /organism="Eptaretus burgeri"
      /db_xref="taxon:7764"
      /sex="female"
      /dev_stage="adult"
      <1..1063
      /codon_start=2
      /product="EphA"
      /protein_id="BA84731.1"
      /db_xref="GI:6002425"
      /translation="EVCSGYLKLGRREVCAIKTKAGTYEQORDFLCASIMGCF
      DHNITIMEGVITKSKPVMITPEFMNGSLDLFRKNDGQFTAMOLVGMRLGIASGK
      YLADMSYVHRDLAARNILVNGNIVCKVSDGSLSRVLEDDDAAYRGGKIPIRMTAP
      EATAFRKFTSASDVMSYGIYVMEVITSGERPYWMSNOVDYIKAEIGYRLPAPDCPL

```

```

ALHQLMDCCORERAVRPKEGHITVTLIDKIRNPATLKPVANEGTMRVPGSLDQSTD
NPATRVPEMLFAIKMGRYADNNAAGYTTLLDSVAQLDAGLVIGITLIGHOKKILS
NIOTMKAEMIDLHSRGVH"

BASE COUNT      389 a      371 c      421 g      382 t

Alignment Scores:
Pred. No.:      0.202      Length:      1563
Score:          12.00      Matches:      12
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    1.09%      Indels:      0
DB:             5      Gaps:      0

US-09-397-967-16 (1-1099) x AB025541 (1-1563)

QY      941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
      |||||||
Db      329 GTACACCGTGACCTTGCTGCCAGAAATATCTGCTT 364

RESULT 88
LOCUS      HSURFRET      1568 bp      mRNA      linear      PRI 14-JAN-1992
DEFINITION "H.sapiens urf-ret mRNA.
ACCESSION      X56348
VERSION      X56348.1      GI:37611
KEYWORDS      urf-ret mRNA.
SOURCE      Homo sapiens.
ORGANISM      Homo sapiens.
      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
      Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE      1 (bases 1 to 1568)
      Kunieda,T.
      Direct Submission
      Submitted (03-SEP-1990) T. Kunieda, IMAMICHI INSTITUTE FOR ANIMAL
      REPRODUCTION, FUKAYA 1103, DEJIMA-MURA, NIINHARI-GUN, IBARAKI
      300-01, JAPAN
      2 (bases 1 to 1568)
      Nomura,N. and Ishizaki,R.
      Cloning of an activated human ret gene with a novel 5' sequence
      fused by DNA rearrangement
      JOURNAL      Gene 107 (2), 323-328 (1991)
      MEDLINE      92084126
      PUBMED      1748302
      location/Qualifiers
      1.1568
      /organism="Homo sapiens"
      /isolate="H116"
      /db_xref="taxon:9606"
      /cell_line="NIH3T3 cells transformed by human DNA"
      /dev_stage="stomach cancer"
      194..1501
      /gene="urf-ret"
      194..1501
      /gene="urf-ret"
      /codon_start=1
      /protein_id="CAA39792.1"
      /db_xref="GI:37612"
      /translation="MEELIVELRLFLFEILDHEVLTSTVREKKAVITNILLRQSSGCF
      DYVDHAKQKQETASLPAPQMPPEIPOPMLDPPKKEFPKKNIVLKGITIGEGEFGVY
      KATAFHLKGRAGTIVAVKMLKENASPELRDLSEFNVLKOVNHHVITKLGYACSD
      GPTLLIVFAKYGLSGFLRESKRVGPYLGSGSSNSLSDHPDERALTMGLISFA
      MQISQCMOYLAEMKILVHRDLAARNILVAERKKKISDFGLSRDYVEDSVYKRSQRI
      PVKMAIESLFDEHITYGSDVMSFGVLMEIVTIGNPPYGLIPERLFINLTKGSHME
      RPDNCSEEMVRLMLQCMKQKQEPVAFADISKDLKEMVKKRDRYLDLAASTPDSLLY
      DGLSDEEPPIVVCNNAIPRALPSWIENKLYGRISHAFTR"

BASE COUNT      359 a      430 c      451 g      328 t

Alignment Scores:
Pred. No.:      0.202      Length:      1568
Score:          12.00      Matches:      12

```

Percent Similarity: 100.00%
 Best Local Similarity: 100.00%
 Query Match: 1.09%
 DB: 9
 Gaps: 0

US-09-397-967-16 (1-1099) x HSURERET (1-1566)

OY 941 VALHISARGASPLEUALAALARGASNIILEUVAL 952
 DB 893 GTTCATCGGACACTTGGCAGCAGAACATCTGTGTA 928

RESULT 89
 AB025538
 LOCUS Eptatretus burgeri mRNA for Ephb, partial cds. 1582 bp mRNA linear VRT 18-NOV-1999
 ACCESSION AB025538
 VERSION AB025538.1 GI:6002418
 KEYWORDS Ephb.
 SOURCE Eptatretus burgeri adult female cDNA to mRNA.
 ORGANISM Eukaryota; Metazoa; Chordata; Gracilata; Hyperotretii; Myxiniiformes;
 Myxiniidae; Eptatretinae; Eptatretus.

REFERENCE 1 (sites)
 Suga, H., Hoshiyama, D., Kuraku, S., Katoh, K., Kubokawa, K. and Miyata, T.

TITLE Protein tyrosine kinase CDNA from amphioxus, hagfish, and lamprey: isoform duplications around the divergence of cyclostomes and gnathostomes

JOURNAL J. Mol. Evol. 49 (5), 601-608 (1999)

MEDLINE 20020330
 REFERENCE 2 (bases 1 to 1582)

AUTHORS Suga, H.
 JOURNAL Direct Submission
 Submitted (31-MAR-1999) Hiroshi Suga, Faculty of Science, Kyoto University, Department of Biophysics, Kitashirakawa-Oiwakecho, Sakyo-ku, Kyoto 606-8502, Japan
 (E-mail:suga@biophys.kyoto-u.ac.jp, Tel:+81-75-753-4224, Fax:+81-75-753-4223)

FEATURES
 source location/Qualifiers

1..1582
 /organism="Eptatretus burgeri"
 /db_xref="taxon:7764"
 /sex="female"
 /dev_stage="adult"
 <1..1063
 /codon_start=2
 /product="Ephb"
 /protein_id="BA84728.1"
 /db_xref="GI:6002419"
 /translation="EVCSGRLLKPKREMSVAITLAKGYTDRODRFLSEASTMGOF
 DHPNVILHEGVYTKSRPYMITTEEMNGSILDSFLRNDGQPTVTOYGMRLGRTAGMK
 YLADMSYVRLDLARNTILVNSLICYKSDRELSFLLEDTSPTITISAMGKIPINW
 APEAIQYRKPTTASDWSYCYTMEVMSYGERPIYDMSNDQVNAIADYRLPAPMOC
 PALHQLMLDSWOKERARERKEGIVNTLKLITPALSILTVACVPGISQPLDRLT
 PVTSTFVTDMDALIKGRYKRENNAGSSFDVYSRMTAEDILRIIGVILAGHOKKI
 LNSTQTMKVNQNOIOTVOV"

BASE COUNT 422 a 348 c 427 g 385 t
 ORIGIN

Alignment Scores:

Pred. No.: 0.204 Length: 1582
 Score: 12.00 Matches: 12
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.09% Indels: 0
 Gaps: 0

US-09-397-967-16 (1-1099) x AB025538 (1-1582)

OY 941 VALHISARGASPLEUALAALARGASNIILEUVAL 952
 DB 329 GTTCATCGGACACTTGGCAGCAGAACATCTGTGTA 364

RESULT 90
 AB049595
 LOCUS Ephydatia fluviatilis EFPTK150 mRNA for protein tyrosine kinase, complete cds. 1585 bp mRNA linear INV 27-DEC-2001
 DEFINITION AB049595.1 GI:18146651
 ACCESSION AB049595
 VERSION AB049595.1
 KEYWORDS Ephydatia fluviatilis cDNA to mRNA.
 SOURCE Eukaryota; Metazoa; Porifera; Demospongiae; Ceractinomorpha; Haplosclerida; Spongillidae; Ephydatia.

REFERENCE 1
 Suga, H., Katoh, K. and Miyata, T.

TITLE Sponge homologs of vertebrate protein tyrosine kinases and frequent domain shufflings in the early evolution of animals before the parazoan-eumetazoan split

JOURNAL Gene 280 (1-2), 195-201 (2001)

MEDLINE 21601119
 REFERENCE 2 (bases 1 to 1585)

AUTHORS Suga, H.
 JOURNAL Direct Submission
 Submitted (03-OCT-2000) Hiroshi Suga, Dept. Biophys., Grad. Sch. Sci., Kyoto Univ.; Kitashirakawa-Oiwakecho, Sakyo-ku, Kyoto, Kyoto 606-8502, Japan (E-mail:suga@biophys.kyoto-u.ac.jp, Tel:81-75-753-4224, Fax:81-75-753-4223)

FEATURES
 source location/Qualifiers

1..1585
 /organism="Ephydatia fluviatilis"
 /db_xref="taxon:31330"
 1..1585
 /gene="EFPTK150"
 8..1420
 /gene="EFPTK150"
 /codon_start=1
 /product="protein tyrosine kinase"
 /protein_id="BA82423.1"
 /db_xref="GI:18146652"

/translation="MGSCSSSPKQDNTNYKTELKDKGVQVEHYPIYCKDYDSRT
 DDPLGFKGDLMIYISREYDWMFAKLDKTRKGYIPSTVAEYKSLDAEWFGLQIK
 RIAERILNPPVAKVESFLIRDSNPQDSSLVKQDQRYRVRRLDGLFLVNG
 VWFQSLDVLVHYRTQKQDGLCNLCPCLQTEKPPQAGLSRQANKMEIDKTOIOLKT
 KLQAGMLGEVMEGVMNGTTSVAVKTRPKSVSEVFLQARIVKLRHRLIOLVAVC
 TKEEPIYVMELEMKYGLVEYLKGEGRTLKIERLVDAIOVASGMSYLRQONYIHCDL
 AARNILVEGHGCKVADIGLTKYIVDKMNEAGCAKRPKWTAPAIIVRPFSTKSYW
 SFGVLIYEITTYGRPYVPGITDHEVLEKIQGYRMCPFCRCREYITDMLDCMHEDPA
 SRPTFTLOMLEPFRNSG"

BASE COUNT 441 a 322 c 418 g 404 t
 ORIGIN

Alignment Scores:

Pred. No.: 0.204 Length: 1585
 Score: 12.00 Matches: 12
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.09% Indels: 0
 Gaps: 0

US-09-397-967-16 (1-1099) x AB049595 (1-1585)

OY 1003 SerAspValTrpSerPheGlyValLeuTYrcLu 1014
 DB 1172 TCGGATGTATGTCATTTGGGTTGCTTTATGAG 1207

RESULT 91
 MM018084

LOCUS MMU18084 1602 bp mRNA linear ROD 02-FEB-1996
 DEFINITION Mus musculus receptor tyrosine kinase mRNA, partial cds.
 ACCESSION U18084
 VERSION U18084.1 GI:1173482

KEYWORDS mouse.
 SOURCE Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. Lickliter, J.D., Smith, F.M., Olsson, J.E., Mackwell, K.L. and Boyd, A.W.

TITLE Embryonic stem cells express multiple Eph-subfamily receptor tyrosine kinases

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 93 (1), 145-150 (1996)

MEDLINE 96133894

PUBMED 8552593

REFERENCE 2 (bases 1 to 1204)

AUTHORS Lickliter, J.

TITLE Direct Submission

JOURNAL Submitted (04-DEC-1994) Jason Lickliter, Lions Laboratory, Walter and Eliza Hall Institute of Medical Research, Royal Parade, Parkville, Victoria, 3050, Australia

COMMENT On Feb 2, 1996 this sequence version replaced gi:1172083.

FEATURES Location/Qualifiers

source 1. .1602

/organism="Mus musculus"

/strain="129/SV"

/db_xref="taxon:10090"

/clone="35C15"

/cell_line="W9.5"

/dev_stage="embryonic stem cells"

<1. .>1602

/codon_start=-2

/product="receptor tyrosine kinase"

/protein_id="AAC52384.1"

/db_xref="GI:1173483"

/translation="CPTGFYRVDMNTLRCLKCPQHSIAESGSLTCTCENGHYRARGE GPVACNRPSPASGLTSTQVGLSEPTANTFTVKQNRNGLSSSPSSALS PCDCGCGVFRSPAAGTSTQVGLSEPTANTFTVKQNRNGLSSSPSSALS INMHASLSGLSLKLVKRPRLQELTMASRPNGNCTYLAVINDOEHWL EPRVLRQLQPDYIVRTRVLAPLGPPSPHREFTSPYRSKLTGRTIVTGL LIGTALIGIVFRSROQRORQREKRTNGREDKMLKPIVDLQATDEPAQCAL DFAQLDPAVLIVDTVIGEGYEGYRGALRLSDCKTVAIKTLRDTSDGYWNNL REATMGOFNPHILRLGIVITRKRPIMITEEMNGALDAFLKREGQLAPQOLVAM LIGTASGMNCSGNYVRDLAARNILNOLCKVSDLGTLRLDDFDCTYETOGCK IPIRWTPAGALAHRIFTTASDVMSFG"

BASE COUNT 391 a 442 c 439 g 330 t

ORIGIN

Alignment Scores:

Pred. No.: 0.206 **Length:** 1602

Score: 12.00 **Matches:** 12

Percent Similarity: 100.00% **Conservative:** 0

Best Local Similarity: 100.00% **Mismatches:** 0

Query Match: 1.09% **Indels:** 0

DB: 10 **Gaps:** 0

US-09-397-967-16 (1-1099) x MMU18084 (1-1602)

Oy 941 valhisargaspleualalalarqanlleuval 952

|||||

Db 1400 gtccatagagacctgcgcagacacattgtggtc 1435

RESULT 92

HUMERKIP 1618 bp mRNA linear PRI 29-MAY-2002

LOCUS Homo sapiens mRNA for large erk/ceks tyrosine kinase, partial cds.

DEFINITION D37827

ACCESSION D37827.1 GI:1060894

VERSION large erk/ceks tyrosine kinase.

KEYWORDS Homo sapiens 4-month fetal brain cDNA to mRNA, clone_lib:1ambda

SOURCE 910.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1

AUTHORS Chan, J. and Watt, V.M.

TITLE eek and erk, new members of the eph subclass of receptor protein-tyrosine kinases

JOURNAL Oncogene 6 (6), 1057-1061 (1991)

MEDLINE 91296384

REFERENCE 2

AUTHORS Iwase, T., Tanaka, M., Suzuki, M., Naifo, Y., Sugimura, H. and Kino, I.

TITLE Identification of protein-tyrosine kinase genes preferentially expressed in embryo stomach and gastric cancer

JOURNAL Biochem. Res. Commun. 194 (2), 698-705 (1993)

MEDLINE 93343925

PUBMED 83343925

REFERENCE 3 (bases 1 to 1618)

AUTHORS Saito, T., Naohiko, S., Kitahara, M., Murata, M., Yamamoto, Y., Hori, T. and Matsuda, Y.

TITLE Identification of human erk gene as a putative receptor tyrosine kinase and its chromosomal localization to 1p36.1: a comparative mapping of human, mouse and rat chromosomes

JOURNAL 4 (bases 1 to 1618)

REFERENCE Saito, T.

AUTHORS Direct Submission

TITLE Submitted (01-AUG-1994) Toshiyuki Saito, National Institute of Radiological Sciences; 9-1 Anagawa 4-chome, Inage-ku, Chiba 263, Japan (Tel:043-251-2111(ex.361), Fax:043-256-8301)

COMMENT On Nov 13, 1995 this sequence version replaced gi:1529066.

Submitted (01-Aug-1994) to DDBJ by:

Toshiyuki Saito

National Institute of Radiological Sciences

Division of Genetics

9-1 Anagawa 4-chome, Inage-ku

Chiba 263

Japan

Phone: 043-251-2111 x333

Fax: 043-251-9818.

FEATURES Location/Qualifiers

source 1. .1618

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="1"

/map="1p36.1"

/tissue_type="brain"

/clone_lib="1ambda gt10"

/dev_stage="4-month fetal"

1. .1618

/gene="ERK"

<1. .>1437

/gene="ERK"

/codon_start=1

/product="large erk/ceks tyrosine kinase"

/protein_id="BA07073.1"

/db_xref="GI:1060895"

/translation="RRTVAGYGRYSKMYFOTMEAEVQTSIQEKLPIITSSAAGL VFLIAYVIAIVCNRRGFERADSEYTDKLOHTSGMTGKMTYIDPFTYEDPNAVR EFAKEIDISCVKIEQVIGAGEFGEVCSGLKIDGKEIFVALIKTLKSGTTEYORDEL SEASINGQFDPHPVNIHLEGVYKSTFVMTIITEMENGSLDSEFLRQDQGFYIQVGM LRQIAQMKTLADMYVHDIARNTIVNSNLVCKYSDGSLRFLDDPTNODVJNAIEOD GGRKIPRWTAPEAIIYRKFTSADVMSYGIYVMEVMSYGERPYWMTNODVJNAIEOD YRLPPMDCPAALHQLMDCWQDRNHRKRFGLIVTLKMRNPSLAKMAPLSGII NLPDLRTIPDYTSFNTVDEMLKAIKMGQKESFANAGFTSDVVSQMMEDILRVGV TLGHQKKILINSIQVMRAQMNQIQSVEV"

DB: 411

gene="ERK"

/citation=[2]

/replace="a"

487. .489

gene="ERK"

/citation=[2]

/replace="cgc"

562

gene="ERK"

/citation=[2]

/replace="g"

1570

gene="ERK"

/citation=[2]

/replace="g"

unsure

unsure

unsure

BASE COUNT 376 a 494 c 444 g 304 t
 ORIGIN

Alignment Scores:
 Pred. No.: 0.207 Length: 1618
 Score: 12.00 Matches: 12
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.09% Indels: 0
 DB: 9 Gaps: 0

US-09-397-967-16 (1-1099) x HUMERKIP (1-1618)

OY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
 |||||
 Db 703 GTTCACCGTGACCTGGCTGCCGACACATCTCTGTC 738

RESULT 93
 LOCUS AB049594 1619 bp mRNA linear INV 27-DEC-2001
 DEFINITION Ephedratia fluvialtilis EFPTK45 mRNA for protein tyrosine kinase,
 complete cds.
 ACCESSION AB049594
 VERSION AB049594.1 GI:18146649
 KEYWORDS
 SOURCE Ephedratia fluvialtilis cDNA to mRNA.
 ORGANISM Ephedratia fluvialtilis
 Eukaryota; Metazoa; Porifera; Demospongiae; Ceractinomorpha;
 Haptosclerida; Spongillidae; Ephedratia.

REFERENCE
 AUTHORS 1
 TITLE Suga, H., Katoh, K. and Miyata, T.
 JOURNAL Sponges homologs of vertebrate protein tyrosine kinases and frequent
 domain shufflings in the early evolution of animals before the
 MEDLINE parazoan-eumetazoan split
 REFERENCE Gene 280 (1-2), 195-201 (2001)
 JOURNAL 21601119
 TITLE 2 (bases 1 to 1619)
 AUTHORS Suga, H.
 JOURNAL Direct Submission
 SUBMITTED (03-OCT-2000) Hiroshi Suga, Dept. Biophys., Grad. Sch.
 Sci., Kyoto-Univ., Kitashirakawa-Oiwakecho, Sakyo-ku, Kyoto, Kyoto
 606-8502, Japan (E-mail: suga@biophys.kyoto-u.ac.jp,
 Tel:81-75-753-4224, Fax:81-75-753-4223)

FEATURES
 source location/Qualifiers
 1..1619
 /organism="Ephedratia fluvialtilis"
 /db_xref="taxon:31330"
 1..1619
 /gene="EFPTK45"
 112..1599
 /gene="EFPTK45"
 /codon_start=1
 /product="protein tyrosine kinase"
 /protein_id="BAB82422.1"
 /db_xref="GI:18146650"
 /translation="MGSLSSPDSSDKSKGVEITNCKRQPEPRKARQAPVQV
 VKPYVYGVYDSTDDDLSPKKDMLITITSDGDMWRFRSKOTGEGITPSYVA
 EKSLDAEMLFKIKRYEAKRLIMOSNQISFLIRSEITPDGFSISVDODRVRA
 YRIRLEDGSLFVTRSTFGQHLVEHYKNTDGLCNLLYPCLOAKPOTAGLSRQ
 ANEWEIKQKIKLKLKAGQFGEVWNGTTSVAVTKLPGTMSVEFLQDASI
 MKLHNPQLIOLAVCTKEEPIYIVTELKYGSLLEYLRDGRITLKIDRLADIAVA
 SGMSTLEQOANYIHDLARNTLVGEHGICVADGGLAVDEIEITANTGAKPKIKNT
 APEAMVNRFTIKSDVMSFGVLYEITTYGKRPFGKMTNPVLEKIQOGRMPCPANC
 PKQYHDVLDGWRDPASRPTEFLIOWLEEFENSEGTRDPDALH"

BASE COUNT 449 a 363 c 432 g 375 t
 ORIGIN

Alignment Scores:
 Pred. No.: 0.208 Length: 1619
 Score: 12.00 Matches: 12
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.09% Indels: 0
 DB: 9 Gaps: 0

DB: 3 Gaps: 0

US-09-397-967-16 (1-1099) x AB049594 (1-1619)

OY 1003 SerAspValTrpSerPheGlyValIleuTyrglu 1014
 |||||
 Db 1327 TCAGATGTGTGTCGTCGTTGGGCTTCTTATGAG 1362

RESULT 94
 LOCUS HSM801500 1619 bp mRNA linear PRI 18-FEB-2000
 DEFINITION Homo sapiens mRNA: cDNA DKFZp434C1418 (from clone DKFZp434C1418);
 partial cds.
 ACCESSION AL133666
 VERSION AL133666.1 GI:6599298
 KEYWORDS
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 1619)
 Othenwaelder, B., Obermayer, B., Mewes, H.W., Gassenhuber, J. and
 Wiemann, S.
 TITLE Direct Submission
 JOURNAL Submitted (15-DEC-1999) MIPS, Am Klopfersplitz 18a, D-82152
 Martinsried, GERMANY
 COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
 Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;
 sequenced by Medigenomix (Martinsried/Germany) within the cDNA
 sequencing consortium of the German Genome Project. This clone
 (DKFZp434C1418) is available at the RZPD in Berlin. Please contact
 the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
 Berlin-Charlottenburg, GERMANY; Email: clone@rzd.de Further
 information about the clone and the sequencing project is available
 at <http://www.mips.biochem.mpg.de/proj/cdna/>.

FEATURES
 source location/Qualifiers
 1..1619
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="DKFZp434C1418"
 /tissue_type="testis"
 /clone_lib="434 (synonym: htes3). Vector pSPORT1; host
 DH10B; sites NotI + SalI"
 /dev_stage="adult"
 343..1347
 /gene="DKFZp434C1418"
 343..1347
 /note="Strong similarity to receptor-like tyrosine kinase
 Etk-2"
 /codon_start=1
 /product="hypothetical protein"
 /protein_id="CA863775.1"
 /db_xref="GI:6599298"
 /translation="MKDSPPQVTRKLYLVNEKWFIFASDMAEQQGILVATAAVGS
 FTLVILTFLLIGRCQYIKAKMKEEERNNLQNGHLPFGIKYIPDLYEDPS
 LAVEHFAKEIDPSRIETRIYIGAEFGEVCSGRKTEGKEIPVATIKLGGHNDOR
 RDLFRESIMKQFDPNIRILEGVTRSPPAIGVAFGFCSPFAGFLNIOAPHPYP
 GGSLPRIPRAGRPVIMVYEMNGSISDLRRHDSGHFYIOLYGMIRGASGKRTYS
 DMGVHNDLARNILTVNSLVCKVSDGLSRVLEDDPEAAFTTDLFQTLNLNCSYA"

BASE COUNT 508 a 338 c 362 g 411 t
 ORIGIN

polyA_signal
 polyA_site 1593

Alignment Scores:
 Pred. No.: 0.208 Length: 1619
 Score: 12.00 Matches: 12
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.09% Indels: 0
 DB: 9 Gaps: 0

US-09-397-967-16 (1-1099) x HSM801500 (1-1619)

QY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
 DB 1183 GTTCATCGACGACTGAGCGCTCGAATATCTGCTC 1218

RESULT 95

MMSEK3 1635 bp mRNA linear ROD 27-APR-1995
 LOCUS M.musculus Sek-3 mRNA.
 DEFINITION X76011
 ACCESSION X76011.1 GI:460055
 VERSION
 KEYWORDS Sek gene.
 SOURCE Mus musculus.
 ORGANISM Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

AUTHORS Becker, N., Seltanidou, T., Murphy, P., Mattei, M.G., Topilko, P., Nieto, M.A., Wilkinson, D.G., Charnay, P. and Giliardi-Hebenstreit, P.
 TITLE Several receptor tyrosine kinase genes of the Eph family are segmentally expressed in the developing hindbrain
 JOURNAL Mech. Dev. 47 (1), 3-17 (1994)
 MEDLINE 95034306
 PUBMED 7947319

REFERENCE

AUTHORS Becker, N.
 TITLE Direct Submission
 JOURNAL Submitted (11-OCT-1993) N. Becker, Unite Inserm U368 Ecole Normale, Superieure, Laboratoire de Biologie Moleculaire, du Developpement-Patrick Charnay, U368 INSERM 46 Rue d'Ulm 75005 Paris, FRANCE

FEATURES

SOURCE 1. 1635
 Location/Qualifiers
 /organism="Mus musculus"
 /strain="BALB/c"
 /db_xref="taxon:10090"
 /chromosome="4"
 /map="D.E"
 /clone="15"
 /tissue_type="brain"
 1. 1443
 /gene="Sek-3"
 <1. 1443
 /gene="Sek-3"
 /codon_start=1
 /protein_id="CAA53598.1"
 /db_xref="GI:460056"
 /db_xref="MGI:96611"
 /db_xref="SWISS-PROT:P54763"
 /translation="VRARTVAGYGRYSGMYFQMTTEAYOTSIREKLLPLIVSSAAGLVPIAVVIVAVICNRGRFRADSEYDQLQHTSGHMTGKMYIDIPFYEEDPNEA
 VREFAKEIDISCVKIEQVIGAGEGEVSGHLKPGKREIPIVATIKLSGYTEKORD
 FLSEASIMGQDPHPVTHLHGVVTKSPVPMITEPMESLIDSLRQNDGQFTVLY
 GMLKGIAGMKYTLADMTVYHDLARNTLVNSLYCKSPDGLSRELPDPTSDPTYS
 ALGKIPITRMTPAELIYRKFTSADVSSTGIVMEVSYSEPRPMDTNDVYNAIE
 ODYRLPMPDPSALHQLMDCQKDRNHRKFGQIVNTDKMLRNPSLAKMAPLSS
 GINPLDRIIPDYTSFNTVDLEALIKMGYKESFANAGFTSPDVVSQMMEDILRY
 GVTLAGHOKILINSIQVRAQMNQIQSVEV"
 1. 99
 /gene="Sek-3"
 /note="extra cellular region"
 100. 177
 /gene="Sek-3"
 /note="trans membrane region"
 178. 1443
 /gene="Sek-3"
 /note="intra cellular region"
 BASE COUNT 403 a 474 c 435 g 323 t
 ORIGIN

Alignment Scores:

Pred. No.:

Length:

1635

Score: 12.00
 Percent Similarity: 100.00%
 Best Local Similarity: 100.00%
 Query Match: 1.09%
 DB: 10
 Gaps: 0

US-09-397-967-16 (1-1099) x MMSEK3 (1-1635)

QY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
 DB 709 GTGCACCGTAGCTTGTCTCGAATATCTGCTC 744

RESULT 96

MMU05210 1651 bp mRNA linear ROD 08-JUN-1994
 LOCUS MMU05210
 DEFINITION Mus musculus protein tyrosine kinase Ctk mRNA, complete cds.
 ACCESSION U05210
 VERSION U05210.1 GI:450232
 KEYWORDS
 SOURCE house mouse.
 ORGANISM Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

AUTHORS Klages, S., Adam, D., Class, K., Fargnoli, J., Bolen, J.B. and Penhallow, R.C.
 TITLE Ctk: a protein-tyrosine kinase related to Csk that defines an enzyme family
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 91 (7), 2597-2601 (1994)
 MEDLINE 94195789
 PUBMED 7511815

REFERENCE

AUTHORS Penhallow, R.C.
 TITLE Direct Submission
 JOURNAL Submitted (14-JAN-1994) Robert C. Penhallow, Molecular Biology, Bristol-Myers Squibb Pharmaceutical Research Institute, Route 206 and Provincetown Road, Princeton, NJ 08543-4000, USA

FEATURES

SOURCE 1. 1651
 Location/Qualifiers
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /tissue_type="brain"
 /clone_lib="murine brain, lambda ZAP"
 /dev_stage="adult"
 117. 1514
 /function="protein tyrosine kinase"
 /product="Ctk"
 /codon_start=1
 /protein_id="AA18829.1"
 /db_xref="GI:450233"
 /translation="MPTRMAGTCYQCMKCNRPKGEIAFRKGDWVTLFACEDKSM
 YRANHSSGQGLAAALRHGEALSTDPKLSMPFHKGISQGEALQQLQPEDGFL
 LVRSASHPGGYVLCVSGFRVHYRVLHNDGHLTIDEAVCFCLMKNVHYTKDKKA
 ICTKIVPRRQGAQKSAEELAKAGWLLDQLHLLGAQIGEGAVLQGYLGQKVA
 VKNKICVYTAQAFIDETAVMTKLOHRLVRLGLVILHNGLYIWEHVSQNLVNFRT
 RGRALVSTSLQLOFALHVAEGMELESEKTLVHRDLARNILVSDLVAKYSDGLAKA
 ERKIDSSRPVKKETAPALKNRFSKSDVSGVILWVESYGRAPYPMKSLKESV
 EAVKGRMEPPDCCPSVHTLMSCHEABAPRRPPRIYKELKRELRVGSAPAG
 GQEAEGAPTSQDP"
 BASE COUNT 365 a 445 c 530 g 311 t
 ORIGIN

CDS

Alignment Scores:
 Pred. No.: 0.211
 Score: 12.00
 Percent Similarity: 100.00%
 Best Local Similarity: 100.00%
 Query Match: 1.09%
 DB: 10
 Gaps: 0

US-09-397-967-16 (1-1099) x MMU05210 (1-1651)

QY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952


```

Db 1035 GTGACCGGAGCTGGCTGCTCGAATCTCTGTC 1070
RESULT 97
LOCUS S77473
DEFINITION 1656 bp mRNA linear ROD 25-AUG-1995
ACCESSION S77473
VERSION S77473
KEYWORDS VNK-nonreceptor protein-tyrosine kinase [mice, mRNA Partial, 1656 nt].
SOURCE Mus sp.
ORGANISM Mus sp.
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. Ershler, M.A., Krivtsov, A.V., Krockova, A.V., Bellivskii, A.V. and Visser, J.V., 1999.
AUTHORS
TITLE VNK-a new gene for nonreceptor protein-tyrosine kinase, expressed in the murine brain and hematopoietic system
JOURNAL Dokl. Akad. Nauk. 339 (5), 679-683 (1994)
MEDLINE 95178994
PUBMED 7874011
REMARK Genbank staff at the National Library of Medicine created this entry (NCBI gibs94 165512) from the original journal article.
COMMENT A base 'a' has been added at position 1408 to accommodate the amino acid sequence given.
FEATURES
    source
        1..1656
            /organism="Mus sp."
            /db_xref="taxon:10095"
            /gene="Vnk"
            /note="nonreceptor protein-tyrosine kinase"
            2..1537
            /gene="Vnk"
            /note="nonreceptor protein-tyrosine kinase: this sequence comes from Fig. 1; Protein sequence is in conflict with the conceptual translation; mismatch(421[E->Q])"
            /codon_start=1
            /protein_id="AAB33566.1"
            /db_xref="GI:957296"
            /translation="MRSIASCSELSHWVLDPLAGTLRSGPRKVSRLFGANHPAPAA
            ARPTQRMABPTGCTCCKNSRPRKGEALFRKGDWTTTEACEDSKWRAKHSGSGE
            GLAALALROREALSTDPKLSLMPHFGKISQCEALIQQLPDEDLFTVRESARHGD
            YVLCSFEGRDVTHRYLRHLDHLLTIDAVCFNLDWVEHYTKDGAICTKLVKPRK
            OCAKSEBELAKAGWLDLQHLTLGAQIGEGEFGAVLQGVAKNICKDYTAQ
            ARIDETAVATKQHRNLVRLGLVILHGLYIWEHVSNGNLVNFRTGRALVSTQAQ
            LQPLALHVAEGMEYLESKLVHRDLAARNILVEDLVAKSDFGLAKARKGIDSSRLP
            VKWTAEALKNRFSKSDVMSFGVILMEVFSGYGRAPYPKMSLKREVSRAVERGYMEP
            PDGPGSVHTLMGSCWEAPRRPFRRKIVKELGRDVSVCVSAAGGDEAGSAPTRS
            QDP"
BASE COUNT 355 a 447 c 536 g 317 t 1 others
ORIGIN
Alignment Scores:
    Pred. No.: 0.211 Length: 1656
    Score: 12.00 Matches: 12
    Percent Similarity: 100.00% Conservative: 0
    Best Local Similarity: 100.00% Mismatches: 0
    Query Match: 1.09% Indels: 0
    DB: 10 Gaps: 0
US-09-397-967-16 (1-1099) x S77473 (1-1656)
OY 941 ValHISARGAspleuAlaAlaArgAsnIleLeuVal 952
DB 1061 GTGACCGGAGCTGGCTGCTCGAATCTCTGTC 1096
RESULT 98
LOCUS BC008655
1658 bp mRNA linear PRI 12-JUL-2001

```

```

DEFINITION Homo sapiens, clone IMAGE:3852708, mRNA, partial cds.
ACCESSION BC008655
VERSION BC008655.1 GI:14250439
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. Strausberg, R.
AUTHORS
TITLE Direct Submission
JOURNAL Submitted (25-MAY-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: http://mgc.ncl.nih.gov
COMMENT Contact: MGC help desk
Email: c9apds-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNLI)
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: Villalón, D.K., Luna, R.A., Hale, S.M., Huiy, S., Lu, X., Garcia, A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W., Muzny, D.M., Gibbs, R.A.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNLI at: http://image.llnl.gov
Series: IRAX Plate: 13 Row: a Column: 6
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.
FEATURES
    source
        1..1658
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="IMAGE:3852708"
            /tissue_type="Colon, adenocarcinoma"
            /clone_id="N1H_MGC_65"
            /lab_host="DH10B"
            /note="Vector: pCMV-Sport6"
            <1..763
            /codon_start=2
            /product="Unknown (protein for IMAGE:3852708)"
            /protein_id="AAH08655.1"
            /db_xref="GI:14250440"
            /translation="AAGMKYLANMYVHDLAARNITVNSNLVCKYSDPGLSRVLEDD
            PEATYTSGGKIPITRMVAPPAISYRKFTASDVMSGTYMWEYMTGERPVELSHHE
            VKAINDGFRLLPMDCPAIVLYIQMOCQOERARKPRADIVSLDKIRAPDSIKT
            LADEPVSIRLPSSTGSESVPRIVSEWLESTIKMOQYEHFMAGYTAIEKVVOMTN
            DDIKRGVIRLPGHOKRIYASLGLKDOVTVGIPIT"
BASE COUNT 350 a 488 c 443 g 377 t
ORIGIN
Alignment Scores:
    Pred. No.: 0.212 Length: 1658
    Score: 12.00 Matches: 12
    Percent Similarity: 100.00% Conservative: 0
    Best Local Similarity: 100.00% Mismatches: 0
    Query Match: 1.09% Indels: 0
    DB: 9 Gaps: 0
US-09-397-967-16 (1-1099) x BC008655 (1-1658)
OY 941 ValHISARGAspleuAlaAlaArgAsnIleLeuVal 952
DB 38 GTGACCGGAGCTGGCTGCTCGAATCTCTGTC 73
RESULT 99

```

MUSNTK
LOCUS MUSNTK 1734 bp mRNA linear ROD 01-JUL-1994
DEFINITION Mus musculus tyrosine protein kinase (Ntk) mRNA, complete cds.
ACCESSION L27738
VERSION L27738.1 GI:507289
KEYWORDS tyrosine protein kinase.
SOURCE Mus musculus (strain BALB/c, sub-species domesticus) day 16 fetus thymus cDNA to mRNA.

ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1734)
Chow, L.M., Jarvis, C., Hu, Q., Nye, S.H., Gervais, F.G., Veillette, A. and Melis, L.A.
Ntk: a Csk-related protein-tyrosine kinase expressed in brain and T lymphocytes
Proc. Natl. Acad. Sci. U.S.A. 91 (11), 4975-4979 (1994)

JOURNAL
MEDLINE 94255451
PUBMED 8197166

FEATURES
source
1..1734
/organism="Mus musculus"
/strain="BALB/c"
/sub_species="domesticus"
/db_xref="taxon:10090"
/tissue_type="thymus"
/dev_stage="day 16 fetus"
/gemline
1..1734
/gene="Ntk"
1..97
/note="Ntk"
/note="putative"
98..1615
/gene="Ntk"
/note="amino acid feature: SH2 domain, bp 455..787; amino acid feature: SH3 domain, bp 248..454; amino acid feature: carboxy-terminal tail, bp 1616..1756; amino acid feature: tyrosine kinase domain, bp 788..1525; amino acid feature: unique domain, bp 98..247"
/codon_start=1
/evidence="experimental"
/product="tyrosine protein kinase"
/protein_id="AAB59677.1"
/db_xref="GI:507290"
/translation="MARSSRSVSLAFEGESRDLPVSPRLFGAMHPAPAAAMPTRWAPGTCTKTCENSRPKPELAFKRGDMYTLLEACDKMYRAKHGSGOGLAAALRGREALSTDPKLSMPFHBKISGOEAIQOOLPEDGLFVRESARRHGDVILVSEGRDVIHYRVLHRDGLTIDEAVCEMLDMVEHYTKDKGATCTKLYKPRKGSASAEELAKAGMLIDGLTGAOIGEGEFAVLQGYLGQKVAKNIKODVIAQAFLEBETA VMRLQHRNLVRLGLVILHGLYIVMEHVSAGNLVFLKTRGALVSTQLQFLAHV ABGMEYTESKIKVHRDLAARNILVSEDLVAKVDFGLAKAERGLDSSRLPKWAPAE ALKNGRPSKSDVMSFCVILMEVEFSYGRAPYKMSLKEVSEAVEKGRMEPPDGGGS VHLNMGSCWEAEPAARRPPRKIVYKELGRILRSYGVAAPAGGGAESAVTRSDP"
226..227
misc_difference
/gene="Ntk"
/standard_name="3 nucleotide insert"
/note="approximately 15% of cellular mRNAs contain this insert"
/replace="cag"
1616..1734
/gene="Ntk"
1709..1713
/gene="Ntk"
1734
/gene="Ntk"
polyA_site
BASE COUNT 362 a 480 c 557 g 335 t
ORIGIN

Alignment Scores:
Pred. No.: 0.219 Length: 1734
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x MUSNTK (1-1734)
QY 941 ValHISArgAspLeuAlaAlaArgAsnIleLeuVal 952
Db 1136 GTGCACCGGAGCCTAGCCTGCTCGACATCCTGTC 1171

RESULT 100
S71669 1744 bp mRNA linear PRI 24-OCT-1994
LOCUS lsk-leukocyte carboxyl-terminal src kinase related gene [human,
DEFINITION peripheral T cells, mRNA, 1744 nt].
S71669
S71669.1 GI:559593
ACCESSION
VERSION
KEYWORDS
SOURCE Homo sapiens peripheral T cells.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1744)
McVicar, D.W., Lai, B.K., Lloyd, A., Kawamura, M., Chen, Y.Q., Zhang, X., Staples, J.E., Ortaldo, J.R. and O'Shea, J.J.
TITLE Molecular cloning of lsk, a carboxyl-terminal src kinase (csk) related gene, expressed in leukocytes
JOURNAL Oncogene 9 (7), 2037-2044 (1994)
MEDLINE 94268844
PUBMED 7516063

REMARK GenBank staff at the National Library of Medicine created this entry [NCBI g1dbsg 151068] from the original journal article.
This sequence comes from Fig. 1.

FEATURES
source
1..1744
/organism="Homo sapiens"
/db_xref="taxon:9606"
1..1744
/gene="lsk"
/note="leukocyte carboxyl-terminal src kinase related gene"
125..1522
/gene="lsk"
/note="leukocyte carboxyl-terminal src kinase related gene: This sequence comes from Fig. 1"
/codon_start=1
/protein_id="AAB30995.1"
/db_xref="GI:559594"
/translation="MPTRRMAPGTGCTTCCKEHTRRPKRGELAFKRGVYVTLLECNKS WYRKHNHSGQEGELAAAGALREBRLSADRLSLMPFHBKISGOEAVQLOPPEDGL FLVRESGHPDGYLVCVTFGRDVIHYRVLHRDGLTIDEAVCEMLDMVEHYTKDKGATCTKLYKPRKGSASAEELAKAGMLIDGLTGAOIGEGEFAVLQGYLGQKVAKNIKODVIAQAFLEBETA VMRLQHRNLVRLGLVILHGLYIVMEHVSAGNLVFLKTRGALVSTQLQFLAHV ABGMEYTESKIKVHRDLAARNILVSEDLVAKVDFGLAKAERGLDSSRLPKWAPAE ALKNGRPSKSDVMSFCVILMEVEFSYGRAPYKMSLKEVSEAVEKGRMEPPDGGGS VHLNMGSCWEAEPAARRPPRKIVYKELGRILRSYGVAAPAGGGAESAVTRSDP"
GQADAGSYSPRSOEP"
BASE COUNT 354 a 534 c 582 g 274 t
ORIGIN

Alignment Scores:
Pred. No.: 0.22 Length: 1744
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 9 Gaps: 0

US-09-397-967-16 (1-1099) x S71669 (1-1744)
QY 941 ValHISArgAspLeuAlaAlaArgAsnIleLeuVal 952
Db 1043 GTGCACCGGAGCCTAGCCTGCTCGACATCCTGTC 1078

Tue Apr 29 06:04:25 2003

Search completed: April 28, 2003, 21:08:33
Job time : 7366 secs

us-09-397-967-16.oligo.rge

THIS PAGE BLANK (USPTO)

GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 28, 2003, 13:46:19 ; Search time 6835.27 Seconds

(without alignments)
4679.258 Million cell updates/sec

Title: US-09-397-967-16
Perfect score: 5860
Sequence: 1 MAPSEEPFLIPQRSCSLSS.....RPAFATLSPQLDPLMRGRPG 1099

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 segs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlp
-O/cg2.1/USPTO.spool/US09397967/rnat.26042003.182312.10714/app.query.fasta.1.1486
-DB=GenEmbl -OPM=fastp -SUFFIX=rge -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.coi -LIST=45
-DOCALIGN=200 -THR=SCORE=500 -MINLEN=0 -ALIGN=15 -MODE=LOCAL
-CURRENT=pro NOR=ext -HEAPSIZE=500 -MILEN=0 -MAXLEN=2000000000
-USBR=US09397967@cgn.1.5398-efunal.26042003.182312.10714 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: GenEmbl:
1: gb_da:
2: gb_hlg:
3: gb_in:
4: gb_ov:
5: gb_ov:
6: gb_pat:
7: gb_ph:
8: gb_pl:
9: gb_pr:
10: gb_ro:
11: gb_sts:
12: gb_sy:
13: gb_un:
14: gb_vl:
15: em_da:
16: em_fun:
17: em_hum:
18: em_in:
19: em_mu:
20: em_mu:
21: em_mu:
22: em_ov:
23: em_ov:
24: em_pat:
25: em_ph:
26: em_ro:
27: em_sts:
28: em_un:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5854	99.9	3723	10 MUSPYKIN	L32955 Mouse prote
2	5687.5	97.1	3528	10 MUSJAK3A	L40172 Mus musculu
3	5380.5	91.8	3778	10 RATJAK3	D28508 Rat mRNA fo
4	5223	89.1	4016	10 MUSJAK3H	L33768 Mus musculu
5	4684	79.9	4064	6 AX203081	AX203081 Sequence
6	4604	79.9	4064	9 HSU09607	U09607 Human JAK f
7	4597	78.4	3620	9 HSU31601	U31601 Human tyros
8	4091.5	69.8	218208	2 AC073700	AC073700 Mus muscu
9	4091.5	69.8	223734	2 AC073750	AC073750 Mus muscu
10	3935	67.2	10326	10 NMU71201	U71201 Mus musculu
11	3336	58.9	3644	5 AF034576	AF034576 Gallus ga
12	3228.5	55.1	2914	6 AX203084	AX203084 Sequence
13	2713	46.3	13562	9 HSU70065	U70065 Human JAK3
14	2637.5	45.0	4812	4 AB006011	AB006011 Sus scrofa
15	2636.5	44.9	3435	6 RNU13396	U13396 Rattus norv
16	2629	44.9	3435	6 AF058925	AF058925 Homo sapi
17	2629	44.9	4161	9 AF005216	AF005216 Homo sapi
18	2628	44.8	3500	9 AF001362	AF001362 Homo sapi
19	2621	44.7	5117	6 AX354568	AX354568 Sequence
20	2621	44.7	5117	6 AF058925	AF058925 Homo sapi
21	2619.5	44.7	3629	6 AR136148	AR136148 Sequence
22	2619.5	44.7	3629	6 AR143945	AR143945 Sequence
23	2619.5	44.7	3629	6 I92575	I92575 Sequence 8
24	2613.5	44.6	3629	10 MUSPYK	L16956 Mouse prote
25	2606.5	44.5	4387	5 DRE5690	AJ005690 Danio rer
26	2523.5	43.1	3739	5 AF148993	AF148993 Cyprinus
27	2511	42.8	4042	4 AB036337	AB036337 Sus scrofa
28	2460.5	42.0	2881	9 BC028068	BC028068 Homo sapi
29	2382.5	40.7	41006	9 AC007201	AC007201 Homo sapi
30	2360	40.3	3495	6 AR048200	AR048200 Sequence
31	2360	40.3	3495	6 AR068116	AR068116 Sequence
32	2360	40.3	3495	6 AR071176	AR071176 Sequence
33	2360	40.3	3495	6 AR177639	AR177639 Sequence
34	2360	40.3	3495	6 I62349	I62349 Sequence 2
35	2360	40.3	3495	6 I87981	I87981 Sequence 2
36	2350	40.1	21323	4 AB036336	AB036336 Sus scrofa
37	2239	39.9	3720	5 AF513860	AF513860 Homo sapi
38	1967.5	33.6	3720	5 AF096264	AF096264 Gallus ga
39	1915.5	32.7	5065	4 AB036335	AB036335 Sus scrofa
40	1908.5	32.6	4191	10 S63728	S63728 Mus sp. JAK
41	1904	32.5	3429	6 AR136149	AR136149 Sequence
42	1904	32.5	3429	6 AR143946	AR143946 Sequence
43	1904	32.5	3429	6 I92576	I92576 Sequence 10
44	1904	32.5	3541	6 AX329654	AX329654 Sequence
45	1904	32.5	3541	6 AX409732	AX409732 Sequence

RESULT 1

ALIGNMENTS

MUSPTRYKIN 3723 bp mRNA linear ROD 21-SEP-1994
 LOCUS
 DEFINITION Mouse protein tyrosine kinase (jak3) mRNA, complete cds.
 ACCSSION L32955
 VERSION L32955.1 GI:529238
 KEYWORDS interleukin 2; interleukin 4; Janus kinase; protein tyrosine kinase.
 SOURCE Mus musculus
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 3723)
 REFERENCE Wiltuhn,B.A., Silvenoinen,O., Miura,O., Lai,K.S., Cwik,C., Liu,E.T. and Inhe,J.N.
 TITLE Involvement of the Jak-3 Janus kinase in signalling by interleukins 2 and 4 in lymphoid and myeloid cells
 JOURNAL Nature 370 (6485), 153-157 (1994)
 MEDLINE 94294024
 PUBMED 8022486
 COMMENT This entry has been reported under the accession number L32955 in the Nature article.
 FEATURES
 source location/Qualifiers
 1..3723
 /organism="Mus musculus"
 /strain="BALB/c"
 /sub-species="domesticus"
 /db_xref="taxon:10090"
 /cell_line="Abelson transformed pre-B-cell line 22D6"
 /cell_type="lymphoid and myeloid"
 1..3723
 /gene="jak3"
 1..12
 /gene="jak3"
 13..3312
 /gene="jak3"
 /standard_name="Janus kinase, mjak3"
 /codon_start=1
 /evidence=experimental
 /product="protein tyrosine kinase"
 /protein_id="AA021565.1"
 /db_xref="GI:529239"
 /translation="MAPPSEETPLIPORSGLSSEACALHYLLPPRGRPPQRLSFGDYLAEDLCVRAKAKALIPVHSLPALATEDSCWPPSHIFCIEDVDVLYRLRFEPDMELETCRHRFGKDLTSLIDLHLVLEHFAHRDLSVGRPLVGLSMKEGFEFLAVLDLQMAKEORRGEELKTVSYACAPSLRDVIOQONFVTRRIRRTVYLALPCGRIPGRPYALMAKYLLDLERLHPATTEFRVGLGAEPEGLLEAVDNGIPWSSNDELFOFCDFPEPVDSINAPRVGAGRHRLVTRRMGHIIEAPEGLPEALSFVALVDGYRLLTODSRHFTCKEYAPRRLLEEDVCHGTTITDFAHKAKAGSELPGLYILRRSPDYDSFLTACVQTPGLDPRKCLLRQDSNAPARPCSFSCALQQLS AACWNSGLRVGAAALYLTSCCAPRKRKSNLIYVRGCPAPARPCSFSCALQQLS HTIPDLSLEHNLGHGSFTKIFRGSREVVDETHDEVLLKWDNRHRCMSFLE AASLSQVSYRHLVLLHGVCMAGDSIMVQEFYVGAIDMYLRKHLLASAMKLOVTK OLAYALNTLEKGLPHGNVSAKRYLLAREGGDGNPFIKLSDPGSPVLVLEMLTR IPWAFBECLQENOTICLEADKWKGGATTWEPFORPAITSLERAKKIKKVEQDGLP ALKWTETAGLITQCMADYDGRPSFRALIKDINGLITSDYELSDPTGTGTPSRDEL VAGOLYACODPAFEERHLKYLISLGGNGFSVGLCYRDPGLDGTGLVAVKQDQSH VPDQORDFOREIOLIKALHSDFIYKRVGYSYGRQSLRVMEVYLPSCILDLQHR GLHTRLLFLAMOLCKGMEYLGARCVHDAARNILVESAHVNIAPFGAKLPLIG KDYVVRERPGSPIFWVAPESLSDNIFRSQDSVMSFGVILVELFTYCDKCSPSAEFL RMGPERGSPRLLELLABRRRLPRPTCTEVEQLMQLCMNEPRHDRAFLATLSP QLDPLMRGPRG-
 mat_peptide 13..3309
 /gene="jak3"
 /product="protein tyrosine kinase"
 /standard_name="Janus kinase, mjak3"
 /evidence=experimental
 3313..3723
 /gene="jak3"
 3723
 /gene="jak3"
 polyA_site
 3'UTR
 BASE COUNT 711 a 1162 c 1079 g 771 t
 ORIGIN 7

Alignment Scores:

Pred. No.: 0
 Score: 5854.00
 Percent Similarity: 99.91%
 Best Local Similarity: 99.91%
 Query Match: 99.90%
 DB: 10
 Gaps: 0

US-09-397-967-16 (1-1099) x MUSPTRYKIN (1-3723)

1 MetAlaProSerGIuGIuThrProLeuILeProGlnArgSerSerLeuSer 20
 13 ATGGCACTTCCAACTGAGAGACACCTCTGATCCCTGAGCGCTTGGACCTCTATCC 72
 21 SerGluAlaGlyAlaLeuHisValLeuLeuProProArgGlyProGlyProProGlnArg 40
 73 TCAGAGGAGAGAGCTTGCATGCTCTCTCCGCCGGGAGCTGGCCCTCCACAGCA 132
 41 LeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuGlyValArgAlaAlaGlyAla 60
 133 TTGCATCTCTCTTGGGAGACTTGGCTGAGGATTTATCTGCGAGCTCCAGCGCC 192
 61 CysGlyIleLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSerCys 80
 193 TGTGCTATCTGCTGCTGTTATCATTCGCTGCTGCTGCGCACTGAGCACTTCTTCC 252
 81 TyrPheProProSerHisIlePheCysIleGluAspValAspThrGlnValLeuValTyr 100
 253 TGGTTTCCCAAGACCACTTCTGCATAGAGACGGACACTCAAGCTTGTGCTAC 312
 101 ArgLeuArgPheTyrPheProAspTyrPheGlyLeuGluThrCysHisArgPheGlyLeu 120
 313 AGGCTACCTTTTATTTCCCGACTGTTGGCTGGAGACATGACCGCTTGGGCTG 372
 121 ArgGlyAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLeuPheAlaGln 140
 373 CGCAAGATTGGACCAAGCCATCTTCTACCTACATGTTTAGAACAATCTTTGGCTACG 432
 141 HisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetGlyGluGly 160
 433 CACGCACTGACCTGCTGAGTGGCGGCTCCCGGGGCTTACACATGAAAGACAGGGA 492
 161 GluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluAlaGlnArg 180
 493 GAGTTCGAGCGTGGCGCTGCTGAGCTGGCCAGATGGCTCTGAGCAGGCCAGCGC 552
 181 ProGlyGluLeuLeuGlySerThrValSerTyrIleAspGlyLeuProProSerLeuArgAsp 200
 553 CCAGAGAGAGCTGTAAGACGGTACGTTACAAGGCTGTGCGGCCAGCCTGCGCAT 612
 201 ValIleGlnGlyGlnAspPheValThrArgArgArgIleArgArgThrValValLeuAla 220
 613 GTGATCCAGGGGCAACAACCTGCTGACAGCGAGGCCATCCGAGGACCGTGCTTGGCG 672
 221 LeuLeuProCysGlyArgLeuProGlyArgProTyrAlaLeuMetAlaGlyTyrIleLeu 240
 673 CTGCTGCGCTGGTGGCTGCTGCGCAGCGGACCGGCTGATGCGCAAGATATTTGG 732
 241 AspLeuGluArgLeuHisIleProAlaAlaThrThrGluThrPheArgValGlyLeuProGly 260
 733 GACCTGAGAGCGCTACATCCAGCGGACCAACGAGACCTTCCGTGGGGTCCCGGGC 792
 261 AlaGlnGluGluProGlyLeuLeuArgValAlaGlyAspAspGlyIleProTyrSerSer 280
 793 GCCCAGGAGAGACCGGGGCTTCTGCTGCTGCGGAGGAGAACCGCATCTTGGAGCTCG 852
 281 AsnAspGluLeuPheGlnThrPheCysAspPheProGluIleValAspValSerIleAsn 300
 853 AACGAGAGACTTTCAGACCTTCTGTACTTCCGGAATGCTGAGATGTCAGCATCAAT 912
 301 GlnAlaProArgValGlyProAlaGlyGluHisArgLeuValThrValThrArgMetAsp 320
 913 CAGGCCCCACGCTGTGGGTCCGCGAGGAGACCGGCTGTGCTACACGAGATGAGAC 972

QY 321 GLYHISILELEUGLUAAGLUPHeproGlyLeuProGluAlaLeuSerPheValAlaLeu 340
 DB 973 GGGCAATCTCGTAACGGAGGTTTCGGGGGCTCCGTGAGGCGCTGCTTTCGGGGCCCTC 1032
 QY 341 VALASPGLYTPHeargLeuIleCysAspSerArgHisTPHeargCysLysGluValAla 360
 DB 1033 GTGGATGGGTAATTCGGCGTATCGTGCAGTCAATTAATTTGTGCAAGAGGAGTGGGG 1092
 QY 361 PROProargLeuLeuGluGluGluAlaAspValCysHisGlyProIleThrLeuAspPhe 380
 DB 1093 CCGCCAAAGGTGCTGGAGGAGGAGGCGGAGGTGTCATGAGACCCATCACTAGACTTT 1152
 QY 381 ALAILEHISLysLeuLysAlaAlaGlySerLeuProGlyTHrTHrIleLeuArgArgSer 400
 DB 1153 GGCATCCACAAGCTGAAGGCGCTGGCTCCCTCCAGGACCACTAATTTCTCCGCGCAC 1212
 QY 401 PROGlnAspTyrAspSerPheLeuLeuThrAlaCysValGlnThrProLeuGlyProAsp 420
 DB 1213 CCGCAGGACTATGACAGCTTTCTTTTACCGCGCTCCAGACTCTCTTGGCGCGAC 1272
 QY 421 TYTLyGLYCysLeuIleArgGlnAspProSerGlyAlaPheSerLeuValGlyLeuSer 440
 DB 1273 TACAAAGGCTGCTCATCCGCCAGACCCAGCGGGCTTCTCCCTGGTTGGCTCAGC 1332
 QY 441 GlnProHisArgSerLeuArgLysLeuLeuAlaAlaCysTPHAsnSerGlyLeuArgVal 460
 DB 1333 CAGCCCCACAGAAGCTGGGGAGTGTTCAGAGCTGTGGAAATTTCTGGGCTGGCAGTA 1392
 QY 461 ASPGlyAlaAlaLeuTyrIleThrSerCysGlyAlaProArgProLysGluLysSerAsn 480
 DB 1393 GAGGGGTGGCCCTGACTTAACATCTCTCGCTCCAGACCCCAAGAAAAGTCCAAAT 1452
 QY 481 LEUleValValArgArgGlyCysAsnProAlaProAlaProGlyCysSerProSerCys 500
 DB 1453 TTGATGTGGTGGGAAGGGGTGCAACCCCGGGCTGCGCTGCTCCCGCTCCTC 1512
 QY 501 CysAlaLeuThrGlnLeuSerPheHisThrIleProThrAspSerLeuGluTPHHisGlu 520
 DB 1513 TGTGGCTGTACACAGCTGAGCTTCCACAAATTCACACGAGCAGCTGGAGTGGCAGAG 1572
 QY 521 AsnLeuGlnHisGlySerPheThrLysIlePheArgGlySerArgArgLysValAlaAsp 540
 DB 1573 AACCTGGGTCAGGGTTCTTTTACCAAGATCTTCCTGGCAGAGGGGGAGGTCTGGGAT 1632
 QY 541 GLyGluThrHisAspSerGlyValLeuLeuLysValIleAspSerArgHisArgAsnCys 560
 DB 1633 GGTGAGACACATGACTCGGAAGTCCCTGAAGGTCAATGAGTCCAGACATGGAACTGC 1692
 QY 561 MetGlnSerPheLeuGlnAlaAlaSerLeuMetSerGlnValSerTyrProHisLeuVal 580
 DB 1693 ATGGAGTCTTTTCTGGAACCCGCAAGCTTGATGACCAAGTATCTTACCCTGGACCTGGTG 1752
 QY 581 LEUleuHisGlyValCysMetAlaGlyAspSerIleMetValGlnIlePheValTyrLeu 600
 DB 1753 TTACGTGACAGGGGTCTGCTGCTGGAGACACATCATGTGCTCAGGAATTTTGTATCTTA 1812
 QY 601 GLYAlaIleAspMetTyrLeuArgLysArgLysHisLeuValSerAlaSerTPHysLeu 620
 DB 1813 GGAGCAATTGACATGTACTCGCAAGCGTGGCCACCTGTGTACGCCAGCTGGAACTG 1872
 QY 621 GlnValIleThrLysGlnLeuAlaTyrAlaLeuAsnTyrLeuGluAspLysGlyLeuProHis 640
 DB 1873 CAGGTGACCAAGCAGCTGGCAATGCTCCCTTAACCTTGGAGAGCAAAAGGCTTCTCTCAC 1932
 QY 641 GlysAsnValSerAlaArgLysValLeuLeuAlaArgLysGlyLysAspGlyAsnProPro 660
 DB 1933 GGCACAGTCTCAGCAAGGAGGTGCTCGTGGCTGTGAGGGGGGTGATGGGATCCACT 1992
 QY 661 PHEIleLysLeuSerAspProGlyValSerProThrValLeuSerLeuGluMetLeuThr 680
 DB 1993 TTTCATTAAGCTGATGATCTGTGTGTCAGTCCCACTGTGCTGAGCTGGAAATGCTCAC 2052

QY 681 ASPArgIleProThrValAlaProGlyCysLeuGlnGluAlaGlnThrLeuCysLeuGlu 700
 DB 2053 GACAGAAATACCTGGGTGGCCCCCAATGCTTCAGAGAGGCTCAGACACTGCTTGGAG 2112
 QY 701 ALASPlySTPGLYPHeargLeuAlaThrThrTPHArgLysValAlaPheGlnArgGlyProAlaHis 720
 DB 2113 GCTGACAAAGTGGGCTTTGGAGCCACACAGTGGAGGTTTCCAAAGGGAGCCCCCACC 2172
 QY 721 ILeThrSerLeuGluProAlaLysLysLeuLysPheTyrGluAspGlnGlyIleLeuPro 740
 DB 2173 ATCACTCGCTGTGAGCCCGCCAAAAGCTTGAAGTCTTATGAGACCAAGGACAGCTGGCC 2232
 QY 741 ALAleuLysTPHThrGlnLeuAlaGlyLeuIleThrGlnCysMetAlaTyrAspProGly 760
 DB 2233 GCTTCAATGTGACAGAACTGGGGGACTTATCACACAGTGCATGGCTATGATCTGCTGC 2292
 QY 761 ArgArgProSerPheArgAlaIleLeuArgAspLysAsnGlyLeuIleThrSerAspTyr 780
 DB 2293 CCGCGCCCTCTCTCGAGACTATCTCAAGAGACCTCAAGCGGCTCATTAACATCAATTAC 2352
 QY 781 GlnLeuLeuSerAspProThrProGlyIleProSerProArgAspGluLeuCysValAla 800
 DB 2353 GAGCTCTCTCAGACCCACACACTGCAATCCGAGTCTCTCGAGATGAGCTGTGCTAGCT 2412
 QY 801 GLYAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGlnGluArgHisLeuLysTyr 820
 DB 2413 GCGCGCCACGCTTATGCTGCTCCAGAGACCCCGCATATTCAGAGAGACACCTTAAGTAC 2472
 QY 821 ILeSerLeuLeuGlyLysGlyAsnPheGlySerValGluLeuCysArgTyrAspProLeu 840
 DB 2473 ATCTCTTGTCTGGCAAGGGCAACTTGTGGACGGTGGAGCTGTGCGCTATGACCCCTG 2532
 QY 841 GLYAspAsnThrGlyProLeuValAlaValLysGlnLeuGlnHisSerValProAspGln 860
 DB 2533 GGGGCAATACGGGACCCCTGCTGGAGTGAACAGCTACAGCAGCAGCTGCCAACCAG 2592
 QY 861 GlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuHisSerAspPheIleVal 880
 DB 2593 CAGAGGACTTCCAGCGGGAGATTGAGATCTTAAGGCTCTGACAGGACTTATCTGCTC 2652
 QY 881 LysTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuAlaGluValMetGluTyr 900
 DB 2653 AAGTACCGGGAGTACGCTATGAGGCAAGTGCACAGACCTGCGTTGTGATGAGATAC 2712
 QY 901 LeuProSerGlyCysLeuArgAspLeuLeuGlnArgHisArgGlyLeuHisThrAspArg 920
 DB 2713 TTGCCCAAGGCTGCTGCTGCTGAGACTGTGTCAGGCGCCATCGCGGCTGTACACCGAC 2772
 QY 921 LEUleuLeuPheAlaTPHInIleCysLysGlyMetGluTyrLeuGlyAlaArgArgCys 940
 DB 2773 CTACTGCTGTTCGGTGGCAGATCTGCAAGGGCATGAGTACCTGGGTGGCGCGGTGC 2832
 QY 941 VALHisArgAspLeuAlaAlaArgAsnIleLeuValGlnSerGluAlaHisValLysIle 960
 DB 2833 GTACACCTGACCTGGGCGCGCCGCAACATCTTGTGGAGACCGAGGCCATGATGAAGTTC 2892
 QY 961 ALASPheGlyLeuAlaLysLeuLeuProLeuGlyLysAspTyrTyrValValArgGlu 980
 DB 2893 GCGGACTTCGGGCTGTGCTAACCTGCTGCGCCCTGGAAAGCATACTACGTGTGCTCGCAG 2952
 QY 981 PROGlyGlnSerProIlePheTPHArgLysLeuProGlnSerLeuSerAspAsnIlePheSer 1000
 DB 2953 CCTGGCCAAAGCCCATATCTTGGATGCCCGGAGTCCCTATGCAACATGCTTCTCC 3012
 QY 1001 ArgGlnSerAspValTPHSerPheGlyValValLeuTyrGluLeuPheThrTyrCysAsp 1020
 DB 3013 CGCCATCTGACGTGTGGAGTTTCGGAAGTGTGTGTGACAGGCTCTTCACTACTGCGAC 3072
 QY 1021 LysSerCysSerProSerAlaGluPheLeuArgMetMetGlyProGlyArgGluGlyPro 1040
 DB 3073 AAGAGCTCACCCCATCGCTGAGATCTCTGGCATGATGGGGGCTGAGAGCGTGAAGACCC 3132
 QY 1041 PROLeuCysArgLeuLeuGluLeuLeuAlaGluGlyArgArgLeuProProProThr 1060

```
|||||
Db 3133 CCGCTCGCCGCGCTGCGAGCTGCTGCGAGGCGGACCCCTCCACCACTCCAC 3192
Oy 1061 CysProthrGluValGlnGluLeuMetGlnLeuCysTrpAlaProGluProHisAspArg 1080
Db 3193 TGGCCACCGAGGTTTCAGGAGCTCATGCTGTGGCGGCCGAGCCGACGACCGG 3252
Oy 1081 ProAlaPheAlaThrLeuSerProGlnLeuAspProLeuTrpArgGlyArgProGly 1099
Db 3253 CCAGCTTCGCGACCTGAGCCCGGAGCTGAGCCCGCTGTGGCTGGAAGACCGGA 3309

RESULT 2
LOCUS MUSJAK3A 3528 bp mRNA linear ROD 24-Apr-1996
DEFINITION Mus musculus JAK3 gene, complete cds.
ACCESSION LA0172
VERSION LA0172.1 GI:1019909
KEYWORDS JAK3 gene.
SOURCE Mus musculus (strain BALB/c x 129 F2) (clone: clone-s12-18) neonate
tymus cDNA to mRNA.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 3528)
AUTHORS Gurniak C.B. and Berg L.J.
TITLE Murine JAK3 is preferentially expressed in hematopoietic tissues
and lymphocyte precursor cells
JOURNAL Blood 87 (8), 3151-3160 (1996)
MEDLINE 96184772
PUBMED 8605329
FEATURES
source location/Qualifiers
1..3528
/organism="Mus musculus"
/strain="BALB/c x 129 F2"
/db_xref="taxon:10090"
/clone="clone-s12-18"
/cell_type="T cell precursor"
/tissue_type="thymus"
/dev_stage="neonate"
1..3528
/gene="JAK3"
1..71
/gene="JAK3"
72..3374
/gene="JAK3"
/codon_start=1
/protein_id="AAC42085.1"
/db_xref="GI:101910"
/translaton="MAPPSEETPLIPQSCSLSSSEAGALHVLPPRGCPORLST
FGDYLAEDLCVRAKAGCILPVHSLPALATEDSCFPPSHIFCIDVDQVLYRL
RFPDWFLETCRRAKGLIKDLITSAIDLHLEHLPQHRSDLSGRLPVGLSKEDG
EFLSLAVDLAOMAREOARPGELKTIVSYKALPSLRDIQOENFVTRRIKRTYV
LAIRPVACADRAVLAIRMAKYILDLERLPAATPTPVLGADGAEQGLLRVADNGI
SWSSGDELROTFCDFPEIYDVSIKQAPRGPRAEHLVYTRDGHILKEPGLDE
ALSTVALVDGTFRLICDSRHFEKVEAPRKLLEAEICHPITLDRATIKRAAGSL
PGTYILRRSPQDYSLFLJACVQPIPLGPKGLIKDDPSAEFLVGLSDHRSRL
LAACMNSGLRVGALNLTSCCAPRPEKSLIVRGCTPAPAPRGSPSCALTOIS
FHTIPTDSLEHNEILGHSFTKIPGRREVVDGETHDEVLKVMDSRHNCESL
EASLSMOVSYPHLYLHGVCMAGDSIMQEFYGLADIMYLKRGHLVSMKLOYT
KQALVALNILEDKLPHGNVSARKVLLARBEGDGNPPFIKSDPCVSPVLSLEMLD
RIPWVAPCELOAEOTLCLEADKGFATWVEFSGGPAHITSLPARKLKETECCOL
PALKMTLAGLITQCMAYDPGRSPFRLIRDLNGLITSDYELSLDPGIPSPRDEL
CGGQALACOPAFEEERHLKYISLKGNGSYELCRYPDELNTPGLVAVKOLHS
GPGQDRDFOREIOLKALHSDFIYKRGVSYGPGROSRLVGMTNIPSGCLDFQHRH
ARLFTDRLLEFAMQICGMEYIARCVHRLARNLILVSEAHVKIADDELAKLPL
GKDIYVREPGSPITFWAPESLSDNIFRSQSDWMSGVYLYELFTCDKCSASAE
LRMGPEREGPPLRLLELAEGRRLLPPPTCTPEVLELMOLCMAPSHDRPAFTLS
POLDALWRGPRG"
3'UTR 3375..3528
/gene="JAK3"
polyA_signal 3517..3522
/gene="JAK3"
BASE COUNT 648 a 1104 c 1036 g 740 t
```

```
ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 3528
Score: 5687.50 Matches: 1080
Percent Similarity: 98.19% Conservative: 2
Best Local Similarity: 98.00% Mismatches: 16
Query Match: 97.06% Indels: 5
DB: Gaps: 3

US-09-397-967-16 (1-1099) x MUSJAK3A (1-3528)
Oy 1 MetaIaProProSerGluGluThrProLeuIleProGlnArgSerCysSerLeuSer 20
Db 72 ATGGACACTTCCAAGTAGAGACACCTGTATCCCTAGCGCTTTCACACCTCATTC 131
Oy 21 SerGluAlaGluAlaLeuHisValLeuLeuProProArgGlyProGlyProGlnArg 40
Db 132 TCAGAGCAGAGAGCGCTGCATGTCTCTCCCTCCCGGGAGCTGGGCTCCACGGA 191
Oy 41 LeuSerPheSerPheGlyAspTrpLeuAlaGluAspLeuCysValArgAlaAlaAla 60
Db 192 TTGCATTCCTCTTTGGGAGCTACTGTGAGGATTTATGTGTGCCAGCTCCAGGCC 251
Oy 61 CysGlyIleLeuProValIleHisSerLeuPheAlaLeuAlaThrGluAspPheSerCys 80
Db 252 TGTGGCATCCGCTGTTTATCATTCGCTTTCGCTCGGCACTGAGACTTCTTGGC 311
Oy 81 TrpPheProProSerHisIlePheCysIleGluAspValAspTrpGlnValLeuValTyr 100
Db 312 TGGTTTCCCGCCAGCCACATCTTGCATAGAGAGCGGACACTCAAGCTCTGCTAC 371
Oy 101 ArgLeuArgPheTrpPheProAspTrpPheGlyLeuGluThrCysHisArgPheGlyLeu 120
Db 372 AGGCTACCTTTATTTCCCTGACTGTGTGGCTGGAGACATTCACCGCTTGGGCTG 431
Oy 121 ArgIysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLeuPheAlaGln 140
Db 432 CGCAAGATTTGACCAAGCGCATCTTGACTTACATGTTTAGAACATCTCTTGTGCTAG 491
Oy 141 HisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetIysGluGlnGly 160
Db 492 CACGCACTGACTGTGAGAGTGGGCGCTCCCGGTGGGCTTACATGAGAGACAGGA 551
Oy 161 GluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGlnAlaArg 180
Db 552 GAGTTCCAGACCGCGCGCTGCTGACTTGGCCAGATGAGCTCGTAGACAGGCCAGGCG 611
Oy 181 ProGlyGluLeuLeuIleValSerTyrIleAlaCysLeuProProSerLeuArgAsp 200
Db 612 CCAGAGAGCGCTGMAACGCTCAGTTACAAGCTGTCTGCGGCCAGGCTCGCGGAT 671
Oy 201 ValIleGlnGlyGlnAspPheValThrArgArgIleArgArgIleArgArgIleValLeuAla 220
Db 672 GTGATCCAGGGCCAGAACTTGCTGACACAGGCGCATCCGACAGACCGCTGTGGCG 731
Oy 221 LeuLeuProCysGlyArgLeuProGlyArgPro-TyrAlaLeuMetAlaIleTyrIleLe 240
Db 732 CTGGC-CCGTGTGGCGCTCTCCAGGCGGACCGCTACCGCTCATGGCAAGTATTTCT 790
Oy 240 uAspLeuGluArgLeuHisProAlaAlaIleThrGluThrPheArgValIleGlyLeuProG 260
Db 791 GGACTTGAGGGGCTACATCCAGCGGCGCACACCGAGACCTTCGTTGGGGCTCCCGGG 850
Oy 260 ValGlnGluGluProGlyLeuLeuArgValAlaGlyAspAsnGlyIleProTyrSerSe 280
Db 851 CCGCCAGAGAGAGCGCGGCTTCTCGTGGCGGGGAGACAGGCAATCTCTGGAGCTC 910
Oy 280 rAsnAsp---GluLeuPheGlnThrPheCysAspPheProGluIleValAspValSerI 299
Db 911 CCGGACACGAGAGCTTTTCCAGACCTTCTGTGACTTTCGGAATAATCGTGATGTCAGCAT 970
Oy 299 eAsnGlnAlaProArgValGlyProAlaGlyGluHisArgLeuValThrValThrArgMe 319
```


Db 971 CAAGAGGAGCCCGACGCTGGGGTCCGCGAGGGAGCACCAGGCTGGTCACTGTCACCCAGCAT 1030
Qy 319 TASPGLYHSLILEUAGLUALAgluPheProGlyLeuProGlyUALaleuSerPheValAl 339
Db 1031 GGACGGCCACATCTGGAGAGCGGAGTTCGCGGGCTGGCTGAGGGCTGTCTTCTGGTGGCC 1090
Qy 339 ALeuValASPGLYrPheArgLeuILECysAspSerArgHISrPheCysLysGLIuVa 359
Db 1091 CCTCGTGGATGGGTACTTCCTCCGCTGATCTCGACATCCAGGCACTTATTTCTGCAAGAGGT 1150
Qy 359 LAlaProProArgLeuLeuGLUGLUALAAspValCysHISGLYProILEthLeuAs 379
Db 1151 GGCGCGGCGACGGCTGCTGGAGAGAGGCGGAGCTGTGCATGAGACCATCATCCTTGA 1210
Qy 379 pPheAlaILEHSLILEuLysUALAglYSerLeuProGlyrPheILEuArgAr 399
Db 1211 CTTTGCCATCCACAGACTGAGAGCGGCTGGCTCCCTCCAGGCACTTATTTCTCGCGCG 1270
Qy 399 gSerProGlnAspTYrAspSerPheLeuLeuThrAlaCysValGLInThrProLeuGLYr 419
Db 1271 CAGCCCGCAGAGCATGTGACAGCTTCTTCTTACGCGCTGGCTCCAGACTCCTTGGCGCC 1330
Qy 419 oAspTYrLysGLYCysLeuILEArgGLInAspProSerGLYAlaPheSerLeuValGLYle 439
Db 1331 CGACTACAGAGGCTCCCTCATCCGCGAGAGCCCGAGCGGGCTTCTCCCTGGTTGGCT 1390
Qy 439 uSerGlnProHISArgSerLeuArgGLILEuLeuUALAglYStrPAsSerGLYLeuAr 459
Db 1391 CAGCCAGCCCGCCACAGAGCTCGGGAGTGGCTTCAGCTGCTCGGAATCTGGGGCTGG 1450
Qy 459 gValASPGLYAlaAlaLeuTYrLeuThrSerCysCysAlaProArgProLYSGLIuLysSe 479
Db 1451 AGTAGACGAGTGGCTGCCGTAACCTTAACATCTGGCTGGCTCCAGACCCAGAAAGATC 1510
Qy 479 rAsnLeuILEValAlaArgArgGLYCysAsnProAlaProAlaProGlyCysSerProSe 499
Db 1511 CAATTGTATCTGGTGGCGAAGGGCTGCACCCCGCGCTGCTGGCTGGCTGCCCGGC 1570
Qy 499 rCysCysAlaLeuThrGlnLeuSerPheHISrPheProThrAspSerLeuGLITrPAl 519
Db 1571 CTGCGTGGCGCTGACAGACAGCTAGCTTCACACATTCACAGAGAGAGCTGGATGGGA 1630
Qy 519 sGLInAsnLeuGLYHISGLYSerPheThrLysILEPheArgGLYSerArgArgGLIuValVa 539
Db 1631 CGAGAACCTGGGCTCAAGGCTTCTTACCAAGATCTTCCTGGCGCCAGCGGAGAGTGCCT 1690
Qy 539 LAspGLYGLUThrHISAspSerGLIuValLeuLeuLysValLMeLAspSerArgHISArgAs 559
Db 1691 GCATGGTAGACACATGACTCGGAAGTCTCTGGAAGTCATGAGACTCCAGACATCGGAA 1750
Qy 559 nCysMetGluSerPheLeuGLIUALAAlaSerLeuMetSerGlnValSerTYrProHISle 579
Db 1751 CTGCGATGGAGTCTTTTCTGGAGAGCCGCAAGCTTGATGAGCCCAAGATATCTACCCCACT 1810
Qy 579 uValLeuLeuHISGLYValCysMeLAlaGLYAspSerILEMeValGLInGLuPheValTY 599
Db 1811 GGTGTACTGACACGCGCTGCTGATGGTAGAGACATCATGATGAGGAATTTGTGA 1870
Qy 599 rLeuGLYAlaILEAspMetTYrLeuArgLYsArgGLYHISLeuValISerAlaSerTrPLY 619
Db 1871 TCTAGGAGCAATGTGACATGACTCGTGGCAAGCTGGCCACCTGGGTGAGCCAGCTGGA 1930
Qy 619 sLeuGLInValThrLysGlnLeuAlaTYrAlaLeuAsnTYrLeuGLInAspLYSGLIuLeuP 639
Db 1931 ACTGAGGTGACCAAGACAGCTGGCATATGCTTAACTACTTTGGAGGACAAAGGCTTCC 1990
Qy 639 oHISGLYAsnValSerAlaArgLYsValLeuLeuUALAArgGLUGLYAspGLYAsnPr 659
Db 1991 TCACGGCAACGCTCTCAGACAGGAAGTGTCTCTGCTGAGGGGGTGAATGGAATCC 2050
Qy 659 oProPheILELysLeuSerAspProGLYValISerProThrValLeuSerLeuGLIuMeLe 679

Db 2051 ACCTTTCATTAAAGCTAGATGATCTGTGTCACTGATCCACCTGTGCTAGCCTGGAAATGCT 2110
Qy 679 uThrAspArgILEProTrpValAlaProGLIuCYsLeuGLInGLIuAlaInThrLeuCYsle 699
Db 2111 CAGCGACAGATATACCTGGGTGGGCCCCGGAATGTCTCCAGAGGGGTCAAGACATCTGCTT 2170
Qy 699 uGLIuAlaAspLYStrPGLYpHeGLYAlaThrThrTrpGLIuValIPheGlnArgGLYProAl 719
Db 2171 GGAGGCTGACAAAGTGGGGCTTTGGAGCCACCAACGTTGGAGGTCTCAAGGGGGAGCCGCC 2230
Qy 719 aHISILEThrSerLeuGLUProAlaLysLysLeuLysPheTYrGLInAspGLInGLYle 739
Db 2231 CCACATCACTCTCGCTGGAGCCCGCCAAAAGCTGAAGTCTATGAGGACCAAGGACAGCT 2290
Qy 739 uProAlaLeuLysTrpTrpHISLeuAlaGLYLeuILEPheGlnCysMeLAlaTYrAspPr 759
Db 2291 GCCCGCTCTCAATGAGACAGACTGGCGGACTTATCAACAGTGCATGGCTATGATATCC 2350
Qy 759 oGLYArgArgProSerPheArgAlaILEuArgAspLeuAsnGLYLeuILEThrSerAs 779
Db 2351 TGGCGGGGCGCCCTCTCCGAGCTATCCTCAGAGACCTCAACGGGCTCATTAACATCAGA 2410
Qy 779 pTYrGLILEuLeuSerAspProThrProGLYILEProSerProArgAspGLIuLeuCYsVa 799
Db 2411 TTACAGCTCTCTCAGACCCACACCTGGCATCCCGAGTCTCTCGAGATGAGCTGTGC-- 2468
Qy 799 LAlaGLYAlaGLInLeuTYrAlaCysGLInAspProAlaILEPheGLUGLYuArgHISLeuLY 819
Db 2469 -GGTGGCGCCAGCTCTATGCTGCCAGAGACCCGCCATATTCAGAGAGACACCTTAA 2527
Qy 819 sTYrILESerLeuLeuGLYLYSGLYAsnPheGLYSerValGLIuLeuCYsArgTYrAspPr 839
Db 2528 GTACATCTCTTGTGGCGCAAGGCAACTTTGGACGGTGGAGCTGTGCCCTATGAGACC 2587
Qy 839 oLeuGLYAspAsnThrHISProLeuValAlaValLYSGInLeuGLInHISSerValProAs 859
Db 2588 CCGGGGGAGCAATYAGGGAGCCCTGTGGCACTGAACACCTACACAGCAGCGGGCCAGA 2647
Qy 859 pGLInAlaArgAspPheGlnArgGLIuILEGInLeuLysAlaLysHISerAspPheIL 879
Db 2648 CCAGCAGAGGCACTTCAGCGGGAGATTCAGATCTCTTAAGCTCTGCACAGGACTTCAT 2707
Qy 879 eValLYSTYrArgLYValSerTYrGLYProGLYArgGlnSerLeuArgLeuValMeTGL 899
Db 2708 CCGTCAAGTACCCGGGAGTACAGTATAGGGCAGAGCTGCAGACCTGCGGTGTGATGGA 2767
Qy 899 uTYrLeuProSerGLYCysLeuArgAspLeuLeuInArgHISArgGLY--LeuHISrH 918
Db 2768 GTACCTGCCAGCGGCTGCTGGAGACTTCTTGACAGCGCCATTCGCGCGCTCCACAC 2827
Qy 918 rAspArgLeuLeuPheAlaTrpGLInILECysLYSGIuMeTGLUTrYrLeuGLYAlaAr 938
Db 2828 CGACCGCTACTGCTGTTCGCTTGGCAGATCTGCAGAGGCAATGAGGTACTGGGTGGCGG 2887
Qy 938 gArgCYsValHISArgAspLeuUALAAlaArgAsnILEuValGLIuSerGLIuAlaHISVa 958
Db 2888 CCGTGGCTGACACCGTGAACCTGGCTGGCGGCAAACTCTGGTGGAGAGGAGCGTCAATGT 2947
Qy 958 LLYSILEAlaAspPheGLYLeuAlaLYSLeuLeuProLeuGLYLYsAspTYrTYrValVa 978
Db 2948 GAAAGTCCCGGACTTCGCGCTCGTAACTGCTGCGCCCTGGGAAAGGACTTACTAGTGTGT 3007
Qy 978 LArgGLUProGLYInSerProILEPheTrpTYrAlaProGLIuSerLeuSerAspAsnIL 998
Db 3008 CCGGAGGCTTGCCCAAGCCCATCTTTTGGATATCCCGGAGTGCCTATCTGACAAACAT 3067
Qy 998 ePheSerArgGlnSerAspValITrPSerPheGLYValValLeuTYrGLIuLeuPheThrTY 1018
Db 3068 CTTTCCCGCAATCTGACAGTGTGAGCTTCGAGATGGTGTGTCAGAGCTTCAACCTA 3127
Qy 1018 rCysAspLYsSerCysSerProSerAlaGLIuPheLeuArgMeTMeTGLYProGLIuArgGL 1038
Db 3128 CTGCGACAAGAGCTGCACCCATCCGCTGAGTTCCTTGGCGCATGATGAGGGGCCGAGACCGTGA 3187

QY 261 AlaGlnGluGluProGlyLeuLeuArgValAlaGlyAspAsnGlyIleProThrSerSer 280
 Db 841 GCCCAGAGAGAGCGGGTCCGCTGCTGACAGGGAGATAGGCTCCCGGAGCTCC 900
 QY 281 AsnAsp---GluLeuPheGlnThrPheCysAspPheProGluIleValAspValSerIle 299
 Db 901 AAGGACGACGAGCTTTTCCAGACCTTCTGTGACTTCCCGAATGCTGATCTCAGCATC 960
 QY 300 AsnGlnAlaProArgValGlyProAlaGlyGluHisArgLeuValThrValThrArgMet 319
 Db 961 AAGCAGCTCCACCGCTGGCTCCGAGAGGAGACCGGCTGATGACCATCACCAGATG 1020
 QY 320 AspGlyHisIleLeuGlnAlaGluPheProGlyLeuProGluAlaLeuSerPheValAla 339
 Db 1021 GAGGCTGACATCTGTGAGACCGAGTCCCGGGGCTCCGAGGGCTCTCTTGTGCGG 1080
 QY 340 LeuValAspGlyTyrPheArgLeuIleCysAspSerArgHisTyrPheCysGlyGluVal 359
 Db 1081 CTCGTGATGGTACTTTCGCTGCTGATCTGCGACTCCAGGCAATTCCTTCAGAGAGGTG 1140
 QY 360 AlaProProArgLeuLeuGluGluGlnAlaAspValCysHisGlyProIleThrLeuAsp 379
 Db 1141 GCACCGCCACGCGCTCTGAGAGAGAGCGGAGCTGCTCCATGACCATCAGCTTGAAC 1200
 QY 380 PheAlaIleHisLysLeuLysAlaAlaGlySerLeuProGlyThrTyrIleLeuArgArg 399
 Db 1201 TTTGGCATTCACAGACTGACAGCTGCTGCTCCCTCCGAGCTCTCTTGGGCCCC 1260
 QY 400 SerProGlnAspTyrAspSerPheLeuLeuThrAlaCysValGlnThrProLeuGlyPro 419
 Db 1261 AGCCCGAGGACATGACACCTTCTTACTCTGCTGCTCCAGACTCTCTTGGGCCCC 1320
 QY 420 AspyrLysGlyCysLeuIleArgGlnAspProSerGlyAlaPheSerLeuValGlyLeu 439
 Db 1321 GACTACAGAGGCTCTCTATCCGCAAGACCCAGTGGGCTTCTCCCTCGTGGTCTC 1380
 QY 440 SerGlnProHisArgSerLeuArgGluLeuLeuAlaIleArgTrpAsnSerGlyLeuArg 459
 Db 1381 AGCCAGCGACAGACACCTTACAGAGCTCTCAGACCGCTGAGCATCTGACACGAA 1440
 QY 460 ValAspGlyAlaAlaLeuTyrLeuThrSerCysCysAlaProArgProLysGlyLysSer 479
 Db 1441 GTAGACGCGACTGCTGAACTCCTACATCTTCTGCTCCCGGACCCCAAGGAAAGTCC 1500
 QY 480 AsnLeuIleValValArgArgGlyCysAsnProAlaProAlaProGlyCysSerProSer 499
 Db 1501 AATTGTATCTGCTGCTGAGGCGCCCAACCCACCGCTGCGCGGACCTCCCATCC 1560
 QY 500 CysCysAlaLeuThrGlnLeuSerPheHisThrIleProThrAspSerLeuGlnIleProHis 519
 Db 1561 TGCTGTGCGCTGACTAGTTCAGCTTCCACACATCCAGACAGACGCTGAGTGGCAGC 1620
 QY 520 GlnAsnLeuGlnHisGlySerPheThrLysIlePheArgGlySerArgArgGluValVal 539
 Db 1621 GAGAACTTGTGTACCGCTCTTTTACAGATCTTCATGCGCAAGAGCGGAGGTGTGTG 1680
 QY 540 AspGlyGlnThrHisAspSerGluValLeuLeuLysValMetAspSerArgHisArgAsn 559
 Db 1681 GATGGTGACACATGACACAGAACTCCCTGAGAGTTATGATGATCCAGACACCAAGAC 1740
 QY 560 CysMetGlnSerPheLeuGlnAlaIleSerLeuMetSerGlnValSerTyrProHisLeu 579
 Db 1741 TGCATGAGACTCTTTCTGAGAGCCCAAGCTTGATGAGCAAGATATCTTACCCACCTG 1800
 QY 580 ValLeuLeuHisGlyValCysMetAlaGlyAspSerIleMetValGlnIlePheValTyr 599
 Db 1801 GTGTGTGCTGATGAGTGTGATGCTGAGACACAGCATCATGGTGCAGAGAAATTTGTGTAC 1860
 QY 600 LeuGlyAlaIleAspMetTyrLeuArgLysArgGlyHisLeuValSerAlaSerTyrLys 619
 Db 1861 CTTGAGACAAATTGACACTGCTGCGCAAGGCTGTACACGCTGCGCGCAGCTGAGAA 1920

QY 620 LeuGlnValThrLysGlnLeuAlaTyrAlaLeuAsnTyrLeuGlnAspLysGlyLeuPro 639
 Db 1921 CTCGAGTGACCAAGACGCTGGCTGCTTACCTTACTTGGAGAGCAAGGCCCTCCCT 1980
 QY 640 HisGlnAsnValSerAlaArgLysValLeuLeuAlaArgGlnGlyLysAspLysPro 659
 Db 1981 CATGGCAACGCTCAGAGAGAGAGAGTGTGCTAGCTCGTGAAGGGGTGTATGAGATCCA 2040
 QY 660 ProPheIleLysLeuSerAspProGlyValSerProThrValLeuSerLeuIleLeu 679
 Db 2041 CTTTCTTCAAGCTGAGTATCTGTGCTGACGCCCAACTGTTTGTGAGCTTGAAATGTCTC 2100
 QY 680 ThrAspArgIleProTrpValAlaProGlyCysLeuGlnGluAlaGlnIleThrLeuCysLeu 699
 Db 2101 ACTGACAGATTCCTCGGGTGGCCCGGAGTGTCTCCAGAAAGCTGGAACGCTCAACTTGG 2160
 QY 700 GlnAlaAspLysTrpGlyPheGlyAlaThrThrTrpGluValPheGlnArgGlyProAla 719
 Db 2161 GAGGCTGACAAAGTGGGCTTTGGCGCCACACGCTGGAGGTGTTCAGGGGGCACCCATG 2220
 QY 720 HisIleThrSerLeuGlnProAlaLysLysLeuLysPheTyrGlnAspGlnGlyLeu 739
 Db 2221 CACATTAACCTCGCTGAGCGCCCAAGACCTGAAGTTCTATGAGAGCGGGGACAGCTG 2280
 QY 740 ProAlaLeuLysTrpThrGlnLeuAlaGlyLeuIleThrGlnCysMetAlaTyrAspPro 759
 Db 2281 CCGGCGCTCAAAATGACAGAGCTGAGGGGCTCATCGCACAGTGCATGCGCTATGATCT 2340
 QY 760 GlyArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIleThrSerAsp 779
 Db 2341 GCGCGGCGCCCTCTCTCCGAGACCATCTCAGAAACCTCAAGACGCGCTCATTAACATCAGAT 2400
 QY 780 TyrGlnLeuLeuSerAspProThrProGlyIleProSerProArgAspGluLeuCysVal 799
 Db 2401 TACGAGCTCTCTCAGACCCACACCTGCGCATCCGAAATCCCGAGATGAGCTGTGT--- 2457
 QY 800 AlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGlnGluArgHisLysLeu 819
 Db 2458 GGTGGTGCCAGCTATGCTCTGCGACAGACCCCGCTTCTTCAGAGAGACACCTTAAG 2517
 QY 820 TyrIleSerLeuLeuGlyLysGlyAsnPheGlySerValGlnLeuCysArgTyrAspPro 839
 Db 2518 TACATCTCTCCGCTGCGCAAGGGGAACCTGCGACGCTGAGCTGTGCGCTACGACCT 2577
 QY 840 LeuGlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGlnHisSerValProAsp 859
 Db 2578 CTGGGACACATACGGGAGCCCTGTGTAAGTACAGACAGATTACAGACAGCGGCGCA 2637
 QY 860 GlnGlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuHisSerAspPheIle 879
 Db 2638 CACGAGAGGAGCTTCCAGGAGATTCAGATCTCTAAGGCTCTACACTTTCATC 2697
 QY 880 ValLysTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuArgLeuValMetGlu 899
 Db 2698 GTCAAGTACCGGGGAGTACACTATGGCCGTGGCCAGACGACTCGGGTTAGTGAATGGAG 2757
 QY 900 TyrLeuProSerGlyCysLeuArgAspLeuGlnArgHisArgGly---LeuHisThr 918
 Db 2758 TACCTGCCAGCGGCTGCTGGGGGACTTCTGACAGCGCCACCGCGCGCTGCACAC 2817
 QY 919 AspArgLeuLeuLeuPheAlaTrpGlnIleCysLysGlyMetGluTyrLeuGlnValAlaArg 938
 Db 2818 GAGCGCTACTGCTGTTCCCTGGGAGATCTCAGAGGCGCATGAGTACTGAGCGGACAGC 2877
 QY 939 ArgCysValHisArgAspLeuAlaIleArgAsnIleLeuValGlnSerGlnAlaHisVal 958
 Db 2878 CCTGGTGCATCGTGAACCTGCTGCGCAACATCTGGTGGAGAGTGAAGCGCCATGTG 2937
 QY 959 LysIleAlaAspPheGlyLeuAlaLysLeuLeuPheLeuGlnLysAspTyrTyrValVal 978
 Db 2938 AAGATTGCTGACTTGGCTGCTGCTAAGCTGCTGCTTGGCAAAAGACTACTACGTGTCTC 2997
 QY 979 ArgGluProGlyGlnSerProIlePheThrTyrIleThrProGlnSerAspAsnIle-998

```

Db 2998 CGCGTGGCGGCGCAAGTCCATCTCTGTAGACCCCTGAGTCCCTATCTGACACATC 3057
Oy 999 Phesertrnglinsaraspvaltrpserphegllyvalleuylrcluleupherthy 1018
3058 TTTCTCCGCCCATCTGACGTGTGAGCTTCCGGGTGCTGCTGATGAGCTCTCACCTAC 3117
Oy 1019 CysasplysSerCysSerProserAlagluPhelenuarmetmetglProgluaArglu 1038
3118 AGCAGCAAGACCTCAGCCATCCATGAGTCTGCTGATGATGAGCTCTGAGCCGTGAG 3177
Oy 1039 GlyProProleuGlyArgleuGluLeuLeuAlagluGlyArgArgleuProProPro 1058
3178 GGATCCCGGCTCTGACACCTCTTGTGAGCTGCTGAGAGGCGCCGCTCTCCACACCC 3237
Oy 1059 ProthrCysProthrGluValGlnGluLeuMetGlnLeuGlySTPAlaProgluProHis 1078
3238 TCCACCTCCCTACTGAGCTTCAAGAGCTCATGAGCTGTGCTGTGCTGAGCCACCCACA 3297
Oy 1079 AsparProAlaPheAlaThrLeuSerProGlnLeuAspProleuTyParGlyArgPro 1098
3298 GACCGGCGGCTTTGACACACCTTGAGCCGCCAGCTGATGCGTGTGCGCGGAGAGCC 3357
Oy 1099 Gly 1099
Db 3358 GGA 3360

RESULT 4
MUSJAK3H 4016 bp mRNA linear ROD 13-SEP-1994
LOCUS Mus musculus (clone 32D5-1) protein tyrosine kinase (JAK3) mRNA.
DEFINITION L33768
ACCESSION L33768.1 GI:508533
VERSION 1
KEYWORDS protein tyrosine kinase.
SOURCE Mus musculus cDNA to mRNA.
ORGANISM Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 4016)
Rane,S.G. and Reddy,E.P.
JAK3: a novel JAK kinase associated with terminal differentiation
of hematopoietic cells
Oncogene 9 (8), 2415-2423 (1994)

JOURNAL MEDLINE 94309920
PUBMED 7518579
FEATURES
source
location/Qualifiers
1..4016
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="32D5-1"
/cell_line="32Dc13"
/tissue_lib="lambda gtl1"
7..3906
/note="protein tyrosine kinase"
/codon_start=1
/product="JAK3"
/db_xref="GI:537607"
/translacion="MAPSPEPLIPORSCSLSSSEAGALHVLPPRGPPORLSEF
FGDYLADLTCVRAKACALIPYHSIPALATEPFCMPSPHFLIEDYDVLVRL
RFYFDFGLGCHFRGLRDLTSALIDLHLEHLAORSDIVSGRLVRGIRSMKDG
EFLSLAVDLAQAAREQAPRGLTIVSKACLPRLVITGQGFVRYRRTIRPVYV
LALWVSPARPTATAGOVYLDLERLHAPATETFPVGLGAGAEFGGLRAGDNCS
WSSGDEVILGILRSVYRGAEMRLVKPVATIRKVASPGRADCTTGQGGVNLKVGPG
MEDQGLTWCTRVRLDRGRTRGEGSDMWISGHDTRPVSITGPGSGRREHRLTVTM
GRGGCAGSRKSVITWLSLFLPFNGFCAPRPYVYSITGPGSGRREHRLTVTM
CLSSAVIKGRVPLRCLSLPOLDFOTFCFPRYVYSITGPGSGRREHRLTVTM
DGHILEFPELPELPSFVALVDYFRLICDSRHYRFXEVAARLLEREAECHGPT
LGASGQOMGMAARFVGLWTLWMLPRLDFALHKAASLPGTITLRSQDDSDS
FLTACVOTPLGPDYKGLIROPDSGAFSLVASAPTEACGSCDILFERLVDGAL
NLTSCAPRPREKNLIVRGCCTPAPAGCSPSCALITLSHTITPTDLSLGHENLGH
GSFTKIRPGRRREYVDETHSEVLLVKVMSRHNCESELEAASLMSOVSFHLVLL
HGVCAGDSINWQEVYILGALDMLLRKRGHLVYSWMLQYTKOLAVLNLNLEDKGLPH

```

```

GNSARVILAREGGDGNPEFLKIDPGVSPYLSLEMLDRIIPWAVECLOEAOTLC
LEADKMGFCATTEVEVSGGPAHITTSLEEDOGOLPALMKETLACILITOCMA
LDPGRPSFRPAIILRLDMLITTSYELLQPHILASRLYELGCGPGRHTEER
HLKYISILGKNGSVELCYDPLDMTGLVAVKOLQHSRPOGRFOREIOILALH
SDFIKRVGYSVPGROSRLNMEYLPSCGLRFLDRHRAKIDTDLLEFAMOTCKM
EYIGARCVHRLAARNIILVSEAHVKIADFGIATKLPLGKDYVVRREGGSPINRYA
PESISDNIFFRSQDSVMSFGVLYELFELCDKSCSPAERLRMMGPRESPLCRLEL
LABERRLPPEPCTCEYQELMOLCVAPSRITGQSPAP"
2199..3024
/misc_feature
/note="JAK homology domain II (JH2)"
3076..3876
/misc_feature
/note="JAK homology domain I (JH1)"
BASE COUNT 742 a 1203 c 1235 g 836 t
ORIGIN
Alignment Scores:
Pred. No.: 1,17e-311 Length: 4016
Score: 5223.00 Matches: 1060
Percent Similarity: 80.59% Conservative: 3
Best Local Similarity: 80.36% Mismatches: 32
Query Match: 89.13% Indels: 235
Gaps: 6
Db: 10
US-09-397-967-16 (1-1099) x MUSJAK3H (1-4016)
Oy 1 MetAlaProProSerGlnGluThrProleuIleProGlnArgSerCysSerLeuSerSer 20
Db 7 ATGCACTCCAGTCAAGTGAAGAGACACCTGTGATCCCTGACGCTGTGACGCTTCATCC 66
Oy 21 SerGlnAlaGlyAlaLeuHisValLeuLeuProProArgGlyProGlyProProGlnArg 40
Db 67 TCAGAGCAGAGAGCCCTGATGTGCTCTTCCCTCCCGGAGACCTGAGCCCTCCAGGCA 126
Oy 41 LeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysValAlaArgAlaAlaVal 60
Db 127 TTGCAATTCCTCTTGGGAGCTACTTGGTGAAGATTATGTGTGGAGCTGCCAAGGCC 186
Oy 61 CysGlyIleLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSerCys 80
Db 187 TGTGCGATCGCTGCTGTTTCACTTCCTTTCGCTTTCGCTGCGCAGGAGACTTCTTCTGC 246
Oy 81 TrpPheProSerHisIlePheCysIleGluAspValAspThrGlnValLeuValTyr 100
Db 247 TGGTTTCCCCCAAGCCACATCTTCTGCATGAGAGACCTGACACTCAAGTCTTGCTAC 306
Oy 101 ArgLeuArgPheTyrPheProAspTyrPheGlyLeuGlnThrCysHisArgPheGlyLeu 120
Db 307 AGGCTACGCTTATTTCTCCGACTGCTTGGGCTGGAGACATGTCACCGCTTGGGCTG 366
Oy 121 ArgTyrAspLeuThrSerAlaIleLeuAspLeuHisValLeuGlnHisLeuPheAlaGln 140
Db 367 CCCAAGATTGACCAAGGCGCATCTTGACTTACATGTTTAAACATCTCTTGTGCTAG 426
Oy 141 HisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLysGlnGlnGly 160
Db 427 CACCCAGTGAAGCTGTGAGTGGGCGCTCCCGGTGGGCTTACCATGAAGGACAGGGA 486
Oy 161 GlnPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGlnAlaArg 180
Db 487 GAGTTCCTGAGCTGCGCTGCTGAGCTTGGCCAGATGAGCTCGAGACAGCCACGCG 546
Oy 181 ProGlyGluLeuLeuLysThrValSerTyrLysAlaCysLeuProProSerLeuArgAsp 200
Db 547 CCAGGAGACCTGCTGAAGCGCTGAGTACCAAGGCTGTGCGCCCGAGCTGCGCAT 606
Oy 201 ValIleGlnGlyGlnAspPheValThrArgArgGlyIleArgArgThrValValLeuAla 220
Db 607 GTATCCAGGCGCAGACTTCGTGACACCGCAGCGCATCCGAGGACCGTGTGCGG 666
Oy 221 LeuLeuProCysGlyArgLeuProGlyArgProTyrAlaLeuMetAlaLysTyrIleLeu 240
Db 667 CTGC--CGGTGTGCTGCGCTGCGAGCGCAGCGCTACGCTCATGCGCCAGATATAT-CTG 723

```

```

OY 241 ASPLeuGluArgLeuHisProAlaIleThrPheArgValGlyLeuProGly 260
DB 724 GACGTGGAGCGGCTTACATCCACCGGCCACACCGAGACTTCCGTGGGCTCCCGGGC 783
OY 261 AAlaGlnGluGluProGlyLeuLeuArgValAlaCysAspAsnGlyIleProTrpSerSer 280
DB 744 GCCCGAGGAGGAGCCCGGGCTTCTGGCTGGCGGGGACACCGCATCTCTGGAGCTCC 843
OY 281 AsnAsp----- 282
DB 844 GGGGACACGAGAGTCTGGGCTGGGCTTGGAGAACGGGGTCCGGGGAGCGCGTGGCGG 903
OY 282 ----- 282
DB 904 TTAGTAAGAGAGCCCGTCCCAACATCAGAAAGTGGCTACTCCGGGGCGGCTGATTCC 963
OY 282 ----- 282
DB 964 ACTACGGGGCAGGGGGGGGGGTCATTTAAAGGTGGTCCGGGATGAGCTGCCAG 1023
OY 282 ----- 282
DB 1024 GAGTGAACCTGGGGCGGTGACTAGAAAGGTGAGACTAGATAGGGGGCGGAGCGAGCAA 1083
OY 282 ----- 282
DB 1084 GCGGATAGTATGATGTGATCAGTGGGATGAGCCGACAGAGCCTGTCTTCAGTCCCTTA 1143
OY 282 ----- 282
DB 1144 ACCTCGTCCCCCCCCCCCCCAAGTGAAGTGGAGAGAGAGAGAGAGAGAGCTGTCT 1203
OY 282 ----- 282
DB 1204 GGGTCAGAGTCTGTAATCCCTGGCTCTGTCTCTCTTTTCTTTTCTTTTAAATGGT 1263
OY 282 ----- 282
DB 1264 TTTCGAGACAGAGGTTTCTGTAATAGCTGTGGTCTGGAACTGACTTTGTAGAGCA 1323
OY 282 ----- 282
DB 1324 GGGTGCCTGGAACCTCAGAAATCCGGCTGGCTCTCGAGTCTGTGATTAAGGG 1383
OY 283 ----- 283
DB 1384 CGTGTGCCACTACGCGGCTACTGCTGTCTTCTTCTTCACAGCTTTCCAGACCTTCTGT 1443
OY 290 ASPPheProGluIleValAspValSerIleAsnGlnAlaProArgValGlyProAlaGly 309
DB 1444 GACTTTCGGGAATCTGTGATGTACACATCAGACAGC--CCACGTGTGGTCCGGCAGGG 1501
OY 310 GlnHisArgLeuValIleThrValThrArgMetAspGlyHisIleLeuGlnAlaGluPhePro 329
DB 1502 AG-CACCGGCTGTCTACTGTACACAGAGATGAGCGCCACATCCTGGAAGCGGATTTCCG 1560
OY 330 GlyLeuProGlnAlaLeuSerPheValAlaLeuValAspGlyTyrPheArgLeuIleCys 349
DB 1561 GGGCTCCCTGAGGCGGTCTTCTTCTGCGGCTCTGTGAGATGGTACTTCCGCTGATCTGC 1620
OY 350 ASPSerArgHisTyrPheCysLysGlnValAlaProArgLeuLeuGlnGluGlnAla 369
DB 1621 GACTCCAGGCAATTAATTTTCGAAGGAGGTGGGGCCACGGCTCTGTGAGAGAGAGCG 1680
OY 370 AspValCysHisGlyProIleThr----- 377
DB 1681 GAGCTGTGCCATGAGCCCATCAGTTAGAGAGCTAGTGGGGTCAACAGTGGGGATGGGGA 1740
OY 378 ----- 380
DB 1741 TCGCGAGCAGAGACTGTCTTGGGTTGAGACTGTGCTGTCTTGGCCATGAGTTAGACTTT 1800
OY 381 AlaIleHisLysLeuValAlaIleGlySerLeuProGlyThrTyrIleLeuArgArgSer 400

DB 1801 ----- 1860
OY 401 ProGlnAspTyrIleAspSerPheLeuLeuThrAlaCysValGlnThrProLeuGlyProAsp 420
DB 1861 CCGCAGACATATACACCTTTCTTCTTACCGCTGTGGTCCAGACTCTCTGTGGCCCGGAC 1920
OY 421 TyrLysGlyCysLeuIleArgGlnAspProSerGlyAlaPheSerLeuValGlyLeuSer 440
DB 1921 TACAGGGCTGCTCATTCGCGCAGAGACCAGGGGGCTTTCTCTCTGCTGTCG-CTCAGC 1979
OY 441 GlnProHisArgSerLeuArgGlnLeuAlaAlaCysTrpAsnSerGlyLeuArgVal 460
DB 1980 CAGCCCAAGAGAGCCCTGGGGAGCTGTGGAGC-TGCTGGAATTC-GGGCTGGAGATA 2037
OY 461 AspGlyAlaAlaLeuThrThrSerCysValaProArgProLysGluLysSerAsn 480
DB 2038 GAGGATGCTGCCCTGAACTAACATCTCTGTGGCTCCAGAACCCAAAGAAAGTCCAAAT 2097
OY 481 LeuIleValValArgArgGlyCysAsnProAlaProAlaProGlyCysSerProSerCys 500
DB 2098 TTGATCTGTGGTGAAGGGGCTGACACCCCGGCGCTGGCCCTGCTCCCTCCCTCTGTC 2157
OY 501 CysAlaLeuThrGlnLeuSerPheHisThrIleProThrAspSerLeuGluTrpHisGlu 520
DB 2158 TGTGCCCTACACAGCTGAGTTCACACAAATTCACACGACAGCCTG---GACACAGAG 2214
OY 521 AsnLeuGlyHisGlySerPheThrLysIlePheArgGlySerArgArgGluValAlaAsp 540
DB 2215 AACCTGGTCAAGGTTCTTTACAAAGATCTTCCGTGGCGGAGCGGAGAGGTCTGTGAT 2274
OY 541 GlyGluThrHisAspSerGlnValLeuLeuLysValMetAspSerArgHisArgAsnCys 560
DB 2275 GGTGAGACACATGACTGGAAATCTCTGAAAGTCATGAGTCAACATCAGACATCGAAATGTC 2334
OY 561 MetGluSerPheLeuGlnAlaAlaSerLeuMetSerGlnValSerTyrProHisVal 580
DB 2335 ATGAGATCTTTCTGGAAGCCCGCAAGCTTGATGACCCAAAGTAATCTTACCCAGCTGTG 2394
OY 581 LeuLeuHisGlyValCysMetAlaGlyAspSerIleMetValGlnGluPheValTyrLeu 600
DB 2395 TTACTGCACAGGGCTGTGAGGCTGTGAGACAGCATCATGTGTGAGGAATTTGTGATCTA 2454
OY 601 GlyAlaIleAspMetTyrLeuArgLysArgGlyHisLeuValSerAlaSerTrpLysLeu 620
DB 2455 GGAGCAATTTGACATGTACTGTGGCAAGGCTGGCCACCTGGGTGAGCCAGCTGGAACCTG 2514
OY 621 GlnValThrLysGlnLeuAlaTyrAlaLeuAsnTyrLeuGlnAspLysGlyLeuProHis 640
DB 2515 CAGTGACCAAGACAGCTGCATATGCTTAACTAAGTGGAGGACAAAGGCTTCTCTCAC 2574
OY 641 GlyAsnValSerAlaArgLysValLeuLeuAlaArgGluGlyAspGlyAsnProPro 660
DB 2575 GGCACAGCTCAGCAAGGAGGTCTCTGTGCTGTGGGGGGTGTGGAATCCACCT 2634
OY 661 PheIleLysLeuSerAspProGlyValSerProThrValLeuSerLeuGlnMetLeuThr 680
DB 2635 TTCAATTAAGCTGATGATCTGTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2694
OY 681 AspArgIleProTrpValAlaProGluCysLeuGlnAlaGlnThrLeuCysLeuGlu 700
DB 2695 GACAGAAATACCTGTGGGCGCCCGCAATGTCTCAGAGAGCTCAGACATCTGTGTTGAG 2754
OY 701 AlaAspLysTrpGlyPheGlyAlaThrThrProGluValPheGlnArgGlyProAlaHis 720
DB 2755 GCTACAAAGTGGGCTTTGAGCCACACAGGAGGTGTTCAGCGGGGAGACCCCGCCAC 2814
OY 721 IleThrSerLeuGlnProAlaLysLysLeuLysPheTyrGlnAspGlnGlyGlnLeuPro 740
DB 2815 ATCACCCTGTGTGAGCCCGCAAAAGCTGAAGTTATAGGACACAGGAGACAGCTGCGC 2874
OY 741 AlaLeuLysTrpThrGlnLeuAlaGlyLeuIleThrGlnCysMetAlaTyrAspProGly 760

```

Db 2875 GCTCTCAAAATGACAGAACTGGCGGACTTATCACACAGTGCATGGCGTATGATCTGCGC 2934
 QY 761 ArgArgProSerPheArgAlaIleLeuAArgPheuAsnGlyLeuIleThrSerAspTyr 780
 Db 2935 CGGGGCCCCCTCTCTCCGAGCTATCTCCAGAGACCTCAAGCCCTCATATCAGATTCATAC 2994
 QY 781 GluLeuLeuSerAspProThrProGlyIleProSerProArgAspGluLeuCysValAla 800
 Db 2995 GAGGCTCCCA-GACCCACACCTGGCATCCCGAGTCCCTCAGATGAGCTGTGCGT----- 3049
 QY 801 GlyAlaGluLeuTyrAlaCysGluAspProAlaIlePheGluGluAAsnGlyLeuIleTyr 820
 Db 3050 GGGCCGCGCTATGCTGCGAGACCCGCCATATC-GAGGAGAGACACCTTAAGTAC 3108
 QY 821 IleSerLeuLeuGlyLysGlyAsnPheGlySerValGluLeuCysArgTyrAspProLeu 840
 Db 3109 ATCTCTTGGCGGAGAGGACACTTGGCAGCGTGGAGCTGTGGCGCTATGACCCCTG 3168
 QY 841 GlyAspAsnThrGlyProLeuValAlaValLysGluLeuGlnHisSerValProAspGln 860
 Db 3169 ---GACAAATACGGAGACCCCTGTGGAGTGAACAGCTACAGCACAGCGGCGCACAGCAG 3225
 QY 861 GlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuHisSerAspPheIleVal 880
 Db 3226 CAGAGGAGCTTCCAGCGGAGATTCAGATCCTTAAGGCTGTCACAGGACTTCATCTGTC 3285
 QY 881 LysTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuArgValMetGluTyr 900
 Db 3286 AAGTACCGGGAGTACGATATGGGCGGAGTCCGAGAGCTGCGGTGGTGAAGTAC 3345
 QY 901 LeuProSerGlyCysLeuAArgPheLeuGlnArgHisArg---GlyLeuHisThrAsp 919
 Db 3346 CTGGCCAGGGGCTGCTGGAGAGATTCCTGCGAGCCCATCGCGCGCCCTGCACACCGAC 3405
 QY 920 ArgLeuLeuLeuPheAlaIlePheGlnIleCysLysGlyMetGlyTyrLeuGlyAlaArg 939
 Db 3406 CCGCTACTCTGTGCTGGCAGATCTGCAAGGCGATGAGTACCTGCTGCGCGCGC 3465
 QY 940 CysValHisArgAspLeuAlaIleAsnGlnIleLeuValGluSerGlnAlaHisValLys 959
 Db 3466 TGGTGTACCGCTGACCTGCTGGCGCAACATCTTGGTGAAGCGAGGCTCATGTGAAG 3525
 QY 960 IleAlaAspPheGlyLeuAlaIleLeuLeuProLeuGlyLysAspTyrTyrValAlaArg 979
 Db 3526 ATGCGGAGCTTGGCTCGCTAGAGTGTGCGCCCTGGAAGAGACTACAGGCTGCGC 3585
 QY 980 GluProGlyGlnSerProIlePheTyrTyrAlaProGlySerLeuSerAspAsnIlePhe 999
 Db 3586 GAGCCTGGCCAAAGCCCATCTTGTGTATGCCCCGAGTCCCTATGTGACACATCTTC 3645
 QY 1000 SerArgGlnSerAspValTyrPhePheGlyValValLeuTyrGluLeuPheThrTyrCys 1019
 Db 3646 TCCCGCAATCTACGTGTGAGCTTCGAGTGTGTAGAGCTCTTACCTACTCTG 3705
 QY 1020 AspLysSerCysSerProSerAlaGluPheLeuArgMetMetGlyProGluAlaGlyGly 1039
 Db 3706 GACAAGAGCTGACAGCCATCCGCTGAGTCTGCGATGTGGGCGCTGAGCCTGAAGA 3765
 QY 1040 ProProLeuCysArgLeuGluLeuGluLeuAlaGluGlyArgArgLeuProProPro 1059
 Db 3766 CCCCCGCTGCGCGCTCTGAGAGCTGTGCGAGAGGGCGGAGCGCTCCACACACCTCC 3825
 QY 1060 ThrCysProThrGluValGlnGluLeuMetGlnLeuCysTyrAlaProGluProHisAsp 1079
 Db 3826 ACCGCGCCACCGAGTTCAGAGCTCATGACGCTGTGCGCGCCAG-CCGACAGAC 3884
 QY 1080 ArgProAlaPheAlaThrLeuSerProGlnLeuAspProLeuTyrPargGlyArgPro 1098
 Db 3885 CGGCACACCTTGGACCTGAGGCCGCCAGCTGAGCGCTGTGGCTGGAAGAGCC 3941
 RESULT 5
 AX203081 . AX203081 4064 bp DNA linear PAT 30-AUG-2001
 LOCUS

DEFINITION Sequence 2 from Patent WO0152892.
 ACCESSION AX203081
 VERSION AX203081.1 GI:15392438
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 4064)
 AUTHORS Vastios, G.
 TITLE Jak/stat pathway inhibitors and the uses thereof
 JOURNAL Patent: WO 0152892-A 2 26-JUL-2001;
 GENZYME CORPORATION (US)
 FEATURES
 source 1..4064
 location/Qualifiers
 BASE COUNT 746 a 1292 c 1168 g 858 t
 ORIGIN
 Alignment Scores:
 Pred. NO.: 1.57e-278 Length: 4064
 Score: 4684.00 Matches: 905
 Percent Similarity: 87.61% Conservative: 64
 Best Local Similarity: 81.83% Mismatches: 129
 Query Match: 79.93% Indels: 9
 Gaps: 5
 DB: 6
 US-09-397-967-16 (1-1099) x AX203081 (1-4064)
 QY 1 MetaLapProSerGluGluThrProLeuIleProGlnArgSerCysLeuSer 20
 Db 96 ATGACACCTCCAAATGAGAGAGACGCCCTGATCCCTGAGGCTCATGACCTTGTCC 155
 QY 21 SerGluAlaGlyAlaLeuHisValLeuLeuProProArgGlyProGlyProGlnArg 40
 Db 156 ACGAGAGCTGTGTGCTGCATGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 215
 QY 41 LeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysValAlaArgAlaVal 60
 Db 216 CTATCTTCTCTCTTGGGAGGACCTTGGCTGAGAGACCTGTGCGGAGCTGCGCAAGGCC 275
 QY 61 CysGlyIleLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSerCys 80
 Db 276 AGCGCATCTGCTGCTGTACCACTCTCTTGTGCTGCGCGAGGAGGAGCTGCTGCG 335
 QY 81 TrpPheProSerHisIlePheCysIleGluAspValAspThrGlnValLeuValTyr 100
 Db 336 TGGTCCCGCCGAGCACATCTTCCCTGGAGATGCCACACCCAAAGTCTCTGTAC 395
 QY 101 ArgLeuArgPheTyrPheProAspTyrPheGlyLeuGluThrCysHisArgPheGlyLeu 120
 Db 396 AGGATTCCTTCTTACTCCCAATGTGTTGGCTGGAGAGAGTGCACCGCTTGGGCTA 455
 QY 121 ArgLysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGlnHisLeuPheAlaGln 140
 Db 456 CCGAAGATTGGCCAGCGATCTTCTTACCTGACCGCAGCTGAGACACCTTGTGGCCAG 515
 QY 141 HisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLysGluGlnGly 160
 Db 516 CACCGAGTACCTGTGTGAGTGGGCGCTCCCGTGGGCTCAGTCTCAAGAGCAGAGGT 575
 QY 161 GluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGlnAlaArg 180
 Db 576 GAGTGTCTAGAGCTGCGCTGTGGACCTGCGCGCGAGATGGCGGAGAGAGCCAGCGG 635
 QY 181 ProGlyGluLeuLeuLysThrValSerTyrLysAlaCysLeuProProSerLeuArgAsp 200
 Db 636 CCGGAGAGCTGTGTAAGACTGTCAAGGCTGCTGCTCAAGGCTTGGCGAC 695
 QY 201 ValIleGlnGlyLysAsnPheValThrArgArgArgIleArgArgThrValValLeuAla 220
 Db 696 CTGATCCAGGGCGCTGAGCTTGTGAGCGGAGGAGCTATTTCGAGAGAGGCTGCGAGAGCC 755

OY 221 LeuLeuProCysGlyAArgLeuProGlyAArgPro-TyralaLeuMetAlaIleuTyrlIlele 240
 Db 756 CTGCC-GCGCGTGGCGGCTGCCAGGACGGGACCTGCTATGGCCAAAGTACATCAT 814
 OY 240 uAspleuGluAArgLeuHISProAlaAlaIleThrGluThrPheAArgValGlyLeuProGly 260
 Db 815 GACCTGGAGCGGCTGGATGCCAGCGGGGCGCCAGACCTTCCTGCACCTGGGCGCTCCCTG 874
 OY 260 yAla-----GlnGluGluProGlyLeuLeuAArgValAlaGlyAspAsnGlyIleP 277
 Db 875 GGGCTTGGTGGCCAGCAGCGGCTGGCTGCTCCGGTGGCTGGTACCGCGCATCCG 934
 OY 277 oTrpSerSerAsp---GluLeuPheGlnThrPheCysAspPheProGlyuIleValAs 296
 Db 935 CTGGACCCAGGAGAAACAGAGAGCTCCAGCCCTTCCTGGCACTTCAGAAATGGTAA 994
 OY 296 pValSerIleAsnGlnAlaProAArgValGlyProAlaGlyGluHISArgLeuValThva 316
 Db 995 CATTAGCATCAAGACAGCCCGCGGCTGGCGCGGCGGAGACACCGCTGGTCACTGT 1054
 OY 316 lThArgMetAspGlyHISIleuGluAlaGluPheProGlyLeuProGlyAlaLeuSe 336
 Db 1055 TACCAGACAGACAAACACATTTTAAAGGCCAGATTCCAGGGGCTGCCGAGGCTCTGC 1114
 OY 336 rPheValAlaLeuValAspGlyTyTrPheAArgLeuIleCysAspSerAArgHISArgPheC 356
 Db 1115 GTTCGGTGGCGCTCGAGAGCGGCTTCGCGGTGACACAGACACTCCAGCACTTCCTGTG 1174
 OY 356 sLySGluValAlaProProAArgLeuGluGluGluAlaAspValCysHISGlyProI 376
 Db 1175 CAAGAGGTGGACCGCGCGAGGCTCTGAGAGAGTGGCCAGAGTGGCCAGCGGCCCAT 1234
 OY 376 ePhrLeuAspPheAlaIleHISLyLeuLyAlaAlaGlySerLeuProGlyHISArgP 396
 Db 1235 CACTGTGACCTTGGCATCAACAAGCTCAACAGCTGGGGGCTCAAGCTCTGGCTCATATG 1294
 OY 396 eLeuAArgSerProGlnAspTyTrAspSerPheLeuLeuThrAlaCysValGlnHISPr 416
 Db 1295 TCTCCGGCGGACCGCCAGGACTTGTACAGCTTCCTCCACTGCTGTGTCAGAACCC 1354
 OY 416 oLeuGlyProAspTyTrLySGlyCysLeuIleAArgGluLeuLeuAlaIleCysTrpAsn 436
 Db 1355 CTTGGTCTCGATTATTAAGGGCTGCTCATCGGCGGACCGCCAGGAGAACCTTCCTTCT 1414
 OY 436 uValGlyLeuSerGlnProHISArgSerLeuAArgGluLeuLeuAlaIleCysTrpAsn 456
 Db 1415 GGTGGGCTCAGCGGACCGCCAGCAGCTCTCGAGAGCTCTGGCAACCTGGCTGGATGG 1474
 OY 456 rGlyLeuAArgValAspGlyAlaAlaLeuTyTrLeuThrSerCysAlaIleProAArgPro 476
 Db 1475 GGGGCTGCACGTAGATGGGGTGGCAGTGCCTCACTTCCTGCTATGCCAGACCCA 1534
 OY 476 sGlyLySerAsnLeuIleValAlaAArgArgGlyCysAsnProAlaIleProAlaProGly 496
 Db 1535 AGAAAGTCCAACTGATGTGGTGCACAGAGTGCACAGCCACCCACCATCATCTTGTGT 1594
 OY 496 sSerPro---SerCysCysAlaLeuThrGlnLeuSerPheHISThrIleProHISArgP 515
 Db 1595 TCAGCCCAATCCCAATACACCTGATGATGATGATGATGATGATGATGATGATGATG 1654
 OY 515 rLeuGluTrpHISGluAsnLeuGlyHISGlySerPheThrLySIIlePheAArgGlySer 535
 Db 1655 CTTGGAGTGGCATGAGAACCTGGGCGCATAGGCTCTTCCACAAAGATTACCGGGGCTGTG 1714
 OY 535 gArgGluValAlaAspGlyGluTrpHISAspSerGluValLeuLeuLySIIlePheAsp 555
 Db 1715 CCATGAGTGGCTGATGGGAGGCCGGAAGACAGAGAGTCTCTCTGAAGTCTATGATGC 1774
 OY 555 rArgHISArgAsnCysMetGluSerPheLeuGluAlaAlaSerLeuMetSerGlnValSe 575
 Db 1775 CAAGCACAAAGACTGATGAGTCAATCTCTGGAAGCAGGAGCTGATGAGCAAGTGTCT 1834

OY 575 rTyTrProHISLeuValLeuLeuHISGlyValAlaCysMetAlaGlyAspSerIleMetValG 595
 Db 1835 GTACCGGATCTCTGGTGTCTCCAGCGGCTGGATGCTGGAGAACGCAACCCATGGTGTCA 1894
 OY 595 nGluPheValTyTrLeuGluAlaIleAspMetTyTrLeuAArgLySArgGlyHISLeuValSe 615
 Db 1895 GGAATTTTACACCTGGGGGCGCATAGACATGTATCTGGCAAAAGCTGGCACTGGTGTGC 1954
 OY 615 rAlaSerTyTrLySLeuGlnValThrLySLeuAlaIleTyTrAlaLeuAsnTyTrLeuGln 635
 Db 1955 AGCCAGCTGGAAAGCTGCAGGTGCTCAAAACAGCTGGCTACGCCCTCACTATCTGAGAGA 2014
 OY 635 pLySGlyLeuProHISGlyAsnValSerAlaAArgLySValLeuLeuAlaAArgGlyGly 655
 Db 2015 CAAGGCTGCCCCCATATGCAATGTCTCTCCGGAAGGTGCTCTGGCTCGGGAGCGGGC 2074
 OY 655 yAspGlyAspProPheIleLySLeuSerAspProGlyValAspProThrValLeuSe 675
 Db 2075 TGAATGGAGCCCGCTTCATCAAGCTGATGACCTCGGGGTCAACCCGCTGTGTAA 2134
 OY 675 rLeuGluMetLeuThrAspArgIleProTrpValAlaProGlyCysLeuGlnGluAlaG 695
 Db 2135 CTTGAGATGCTCACCGACAGATCCCTGGGTGGCCCGGAGTGTCTCCGGAGGGCGCA 2194
 OY 695 nThrLeuCysLeuGluAlaAspLySArgLyPheGlyValAlaThrTrpGluValPheG 715
 Db 2195 GACACTTACCTTGGAAAGCTGACAAAGTGGGCTTCGGCCCGCCAGGCTGTGGAAAGTGTAC 2254
 OY 715 nArgGlyProAlaHISIleThrSerLeuGluProAlaLySLeuLySLeuAspPheTyTrGlu 735
 Db 2255 TGGCCTCAACATGCCCATCAAGTGGCTGTGATCTCGTCAAAACTCCAAATTTATAGGA 2314
 OY 735 pGlnGlyLeuLeuProAlaLeuLySArgLyGluLeuAlaGlyLeuIleHISArgPheCys 755
 Db 2315 CCGGAGAGCTGGCGGCGCCAGAGTGGACAGAGCTGGCTCTGATTCAAAGTGCAT 2374
 OY 755 fAlaTyAspProGlyAArgTrpAspSerPheAlaIleLeuAArgAspLeuAsnGly 775
 Db 2375 GGCTATAGAGCGGCTCCAGAGGCTCTCTCCAGGCTGCATTCGAGCTCAATAGCTT 2434
 OY 775 uIleThrSerAspTyTrGlyLeuLeuSerAspProThrProGlyIleProSerProAArg 795
 Db 2435 CATCTTTCAGACTATGAGCTCTTCACACCCCAACCTGTGCTGGCCCTCTGCA 2494
 OY 795 pGluLeuCysValAlaGlyAlaGlnLeuTyTrAlaCysGlnAspProAlaIlePheGlyG 815
 Db 2495 TGGGCTGTGG---AATGGTGGCAGCTCTATGCTGCGCAAGACCCACGATCTTGGAGA 2551
 OY 815 uArgHISLeuLySArgLySerLeuLeuGlyLySArgAsnPheGlySerValGlyLeuCy 835
 Db 2552 GAGACACCTCAAGTACATCTCACAGCTGGGCAAGGCAACTTTTGGCACTGGAGCTGTG 2611
 OY 835 sArgTyTrAspProLeuGlyAspAsnThrGlyProLeuValAlaValLySLeuGlnI 855
 Db 2612 CCGGTATGACCGGCTAGCCCAATACAGAGTGGCTGTGGCCCTGAAACAGCTGCAGCA 2671
 OY 855 sSerValProAspGlnGlnAArgAspPheGlnAArgGlyIleGlnIleLeuLySAlaLeuH 875
 Db 2672 CAGCGGCGCAAGCAGCAGAGGAGCTTCACGGGAGATTCAGATCTCAAACTCTGCA 2731
 OY 875 sSerAspPheIleValLySArgLyGlyValSerTyTrGlyProGlyAArgGlnSerLeuA 895
 Db 2732 CAGTATTTTCAATTCAGAAATGCTGTGTGACCTAAGGCCCGGGCGGCAAGCTGGG 2791
 OY 895 gLeuValMetGlyTyTrLeuProSerGlyCysLeuAArgAspLeuGlnAlaArgHISArg 915
 Db 2792 GCTGTTCATGAGTACCTGCGCCAGGCTCTTGGCGGACTTCCTGCACCGCCACCGCC 2851
 OY 915 y---LeuHISThrAspAArgLeuLeuPheAlaTrpGlnIleCysLySArgMetGly 934
 Db 2852 GCGGCTGTGATGCCAGCGGCTCTTCTTATCTTCGCGAGTCTGCAGAGGCGATGAGAT 2911
 OY 934 rLeuGlyAlaAArgArgCysValHISArgAspLeuAlaAlaArgAsnIleLeuValGlySe 954

Db 636 CCGGAGAGCTGTCAGACTGTCAACAAGCCCTGCTACCCCAAGCCCTGGGAC 695
 Oy 201 ValIeGInGlyGlnAsnPhenValThrArgArgArgIleA-gA-gthValIleuAla 220
 Db 696 CTGATCCAGAGCCCTGAGCTTGTGACGGAGGAGCTATTCGAGAGAGCTCCGACAGCC 755
 Oy 221 LeuLeuProCysGlyArgLeuProGlyArgPro-TyrAlaLeuMetAlaIleuIle 240
 Db 756 CTGGC-GCGCGCTGGCCCTCCAGCGACCGGACCTGCTATGGCCAAAGTACATCAT 814
 Oy 240 uAspLeuGlnArgLeuHisProAlaIleThrGlnIleThrPheArgValGlyLeuProGly 260
 Db 815 GGACCTGGAGCGGCTGGATCCAGCGGGGCGCGGAGACCTTCACACTGGCCCTCCCTCG 874
 Oy 260 yAla-----GInGInGInProGlyLeuLeuArgValAlaGlyAspAsnGlyIlePr 277
 Db 875 GGCCCTGGTGGCCAGCAGCGGCTGGGGCTGCTCCGGCTGGCTGTACCGCGGCATCCG 934
 Oy 277 oTrpSerSerAsnAsp---GluLeuPheGlnIleThrPheCysAspPheProGluIleValAs 296
 Db 935 CTGGACCCAGGAGACAGAGAGGTCTCCAGCCCTTCTGCGACTTTCAGAAATCTGAGA 994
 Oy 296 pValSerIleAsnGlnAlaProArgValGlyProAlaGlyGlnHisArgLeuValThrVa 316
 Db 995 CATTAAGCATCAAGCAGGCGCCCGCGCTGGCGCGCGGAGACACCGCTGGTCACTGT 1054
 Oy 316 lThArgMetAspGlyHisIleLeuGlnAlaGluPheProGlyLeuProGluAlaLeuSe 336
 Db 1055 TACCAGACACACACACCAAGTATTTAGAGCGCGAGTTCCAGGGCTGCCAGAGCTCTGTG 1114
 Oy 336 rPheValAlaLeuValAspArgIleThrPheArgLeuIleCysAspSerArgHisIlePheC 356
 Db 1115 GTTCGTGGCGCTGTGGACGGCTACTTCCGGCTGACACCGAGCTCCAGACTTCTTGTG 1174
 Oy 356 sIysGluValAlaProProArgLeuLeuGlnGlnAlaAspValCysHisGlyProI 376
 Db 1175 CAAGGAGGTGGACCGCCGAGCTGGAGAAAGTGGCGGACAGCAATGGCAGCGGCCCAT 1234
 Oy 376 eThrLeuAspPheAlaIleHisIleuLeuValAlaGlySerLeuProGlyThrTyrI 396
 Db 1235 CACTCTGAGACTTGGCATCAACAAGCTCAAGAGCTGGGGCTCACGTCGGCTCCATGT 1294
 Oy 396 eLeuArgArgSerProGlnAspArgArgAspPheLeuLeuThrAlaCysValAlaIleThrPr 416
 Db 1295 TCTCCGCCCGACGCCCGACAGACTTGTGACAGCTTCCTCCACGTGTGTGCCAAGACC 1354
 Oy 416 oLeuGlyProAspArgIleGlyCysAsnLeuLeuArgGlnAspProSerGlyAlaPheSerI 436
 Db 1355 CCTTGTCTCTGATTAAGGGCTGCTCATCGGGCGACCCACAGAGAACTTCTTCT 1414
 Oy 436 uValGlyLeuSerGlnProHisArgSerLeuArgGluLeuLeuAlaIleCysTrpAsnSe 456
 Db 1415 GGTGGAGCTCAGACCCAGACAGAGCTTGTGAGAGCTCTGGAGACCTGTGGGATGG 1474
 Oy 456 rGlyLeuArgValAspGlyAlaIleLeuThrGlnSerCysAlaProArgProLy 476
 Db 1475 GGGGCTGCACGTAGATGGGTGGTCAGTGACCTCACTTCCTGTATCCCAAGCCCAA 1534
 Oy 476 sGluIysSerAsnLeuIleValAlaArgGlyCysAsnProAlaProAlaProGlyC 496
 Db 1535 AGAAAGTCCAAACCTGATCTGTGGTCCAGAGAGTCCAGAGCCACCAATCATCTTGT 1594
 Oy 496 sSerPro---SerCysAlaLeuThrGlnLeuSerPheHisIleThrIleProHisAsp 515
 Db 1595 TCAGGCCCAATCCCAATACAGCTGAGTCAGATGACATTTCAACAAGATCCCTGTGACAG 1654
 Oy 515 rLeuGlnThrHisGlnAsnLeuGlnHisGlySerPheThrIlePheArgGlySerArg 535
 Db 1655 CCTGAGAGGTGATGAGAACTGGGCGAATGGCTCTTCCAAAGATTTAACCGGGGCTGCG 1714
 Oy 535 gArgGluValAlaAspGlyGlnThrHisAspSerGluValLeuLeuIysValIleMetAsp 555
 Db 1715 CCAATGAGTGGTGGATGGGAGGCGCCAAAGACAGAGGTGCTGCTGAAGGTATGATGTC 1774

Oy 555 rArgHisArgAsnCysMetGlnSerPheLeuGlnAlaIleSerLeuMetSerGlnValSe 575
 Db 1775 CAAGCAACAAGATGATGATGAGTATCTCTGGAGAGGAGAGCTGATGAGCAAGTGTG 1834
 Oy 575 rTyrProHisLeuValLeuLeuHisGlyValCysMetAlaGlyAspSerIleMetValG 595
 Db 1835 GTACCGGATCTCTGCTCTCCAGCGCGTGTGCATAGCTGTGGAGACAGCAATGTTGA 1894
 Oy 595 nGluPheValIleThrLeuGlyAlaIleAspMetIleuArgIleuArgGlyHisLeuValSe 615
 Db 1895 GGAATTTGTACACTGTGGGCGCATAGACATGATGTGCGAAAGCGTGGCCACTGTGCGC 1954
 Oy 615 rAlaSerTrpIleuGlnValThrIleGlnLeuAlaIleAlaLeuAsnTyrLeuGlnAs 635
 Db 1955 AGCCACTGGAGCTGACAGGTGATGATCAACAGCTGGCGCTACGCCCTCAACTATCTGAGA 2014
 Oy 635 pIysGlyLeuProHisGlyAsnValSerAlaArgIleValLeuLeuAlaIleArgGlyG 655
 Db 2015 CAAGGCCCTGCCCATGCTGCAATGTCTGTGCGGAAAGTGTCTGTGGCTGGAGGGGCG 2074
 Oy 655 yAspGlyAsnProProPheIleIleuLeuSerAspProGlyValSerProThrValLeuSe 675
 Db 2075 TGATGGAGACCGCCCTTCATCAAGCTGATGACCTGGGTGACCCCGCTGTGTAG 2134
 Oy 675 rLeuGlnMetLeuThrAspArgIleProTrpValAlaProGlyCysLeuGlnGlnAlaG 695
 Db 2135 CCTGGAGATGCTCACCGAGAGATCCCTGGGTGGCCCGCCAGTGTCTCCGGAGGCGCA 2194
 Oy 695 nThrLeuCysLeuGlnAlaAspIleIleuLeuGlnProAlaIleuLeuIleuLeuPheG 715
 Db 2195 GACACTTAGCTTGAAGGTGACAAAGTGGGCTGTGGGCGCCACGGTCTGGGAAGTGTAG 2254
 Oy 715 nArgGlyProAlaHisIleThrSerLeuGlnProAlaIleuLeuIleuLeuIleuLeu 735
 Db 2255 TGGCTACCAATGCCATCACTGCTGCTGATCTGTGAAGAACTTATATGAGA 2314
 Oy 735 pGlnGlyLeuLeuProAlaIleuLeuIleuLeuIleuLeuIleuLeuIleuLeuIleu 755
 Db 2315 CCGGAGCAGCTGCGCGGCCCGCCCAAGTGGAGACAGCTGCGCTGTGATCAACAGTGCAT 2374
 Oy 755 tAlaIleThrAspProGlyArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLe 775
 Db 2375 GGGCTATGAGCGGGTCCAGAGGCGCTCTTCCGAGCGCTGATGTGCACCTCATATACCT 2434
 Oy 775 uIleThrSerAspIleuLeuLeuSerAspProThrProGlyIleProSerProArgAs 795
 Db 2435 CATCTCTCACTATGAGCTCTCTCAGACCCCAACCTGTGTGCTGTGACCTGTGCA 2494
 Oy 795 pGluLeuCysValAlaGlyAlaGlnLeuIleuIleuIleuIleuIleuIleuIleuIleu 815
 Db 2495 TGGGCTGTGG--AATGTGGCCAGCTGTATGCTGTGCAGAGACCCCAAGCTTTCAGAGA 2551
 Oy 815 uArgHisLeuIleuIleuIleuSerLeuLeuGlyIleuGlyAsnPheGlySerValGluLeu 835
 Db 2552 GAGACACTCAAGATGATCTGTGACAGCTGGGCAAGGGCAACTTGTGGAGCTGTGAG 2611
 Oy 835 sArgTyrAspProLeuGlyAspAsnThrGlyProLeuValAlaValIleGlnLeuGlnI 855
 Db 2612 CCGCTATGACCCGCTACCCCAATACAGTGGCTGTGTGGCGGTGAACAAGCTGTGAGA 2671
 Oy 855 sSerValProAspGlnGlnArgAspPheGlnArgGlnIleGlnIleLeuIleuIleuIleu 875
 Db 2672 CACCGGGCCAGACAGAGAGGAGTTCAGCGGAGATTCAGATCTCAAGACACTGGA 2731
 Oy 875 sSerAspPheIleValIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 895
 Db 2732 CACTGATTTATGTGTAAGATGTGTGTGTGACACTTATGGCCCGGCGGCGAGCTGGG 2791
 Oy 895 gIleuValMetGluIleuLeuProSerGlyCysLeuArgAspLeuLeuGlnArgHisArg 915
 Db 2792 GTGTGATATGAGATACCTGCGGAGGCTGTGGCGGACTTCTGTGACAGGCGACCGCGC 2851

Db 396 AGGATTCCTTTACTTCCCAATTTGGTTGGCTGGAGAAAGTCCACCCCTTCGGGGTA 455
Oy 121 ArgLysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLeuPheAlaGln 140
Db 456 CGCAAGCATTTTGGCCAGGCTATCTTGGACCTGCCAGTCCGGAGCACCTCTTTGGCCAG 515
Oy 141 HisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLysGluGlnGly 160
Db 516 CACCCGACGTGACCTGTGAGTGGGGCCCTCCCTGGGCTTCAGTCTCAAGAGAGAGGT 575
Oy 161 GluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGlnAlaGlnArg 180
Db 576 GAGTGTCTCAGCTGGCCGCTGTGGACCTGGCCCGGATGGCCGAGACAGCGCCAGCGG 635
Oy 181 ProGlyLeuLeuLeuLysThrValSerTyrLysAlaCysLeuProSerLeuArgAsp 200
Db 636 CCGGAGAGCTGCTGAAAGACTGTACACTACAGAGCTGCTCCCTCAAGCTCGCCGAC 695
Oy 201 ValIleGlnGlyLysAsnPheValThrArgArgArgIleArgArgThrValValLeuAla 220
Db 696 CTGATCCAGGGCCTGAGCTTCGTGACGCGAGGCGTATTCGGAGAGCGGTGCCAGAGCC 755
Oy 221 LeuLeuProCysGlyArgLeuProGlyArgPro-TyrAlaLeuMetAlaLysTyrIleLe 240
Db 756 CTGCG-CCCGCTGGCCGCTGCCAGGACGCGGACACTGCTCATGGCCAGTACATCAT 814
Oy 240 uAspLeuGluArgLeuHisProAlaAlaThrThrGluThrPheArgValGlyLeuProGly 260
Db 815 GGCACCTGGAGCGGCTGATCCAGCCGGGCGCCGAGACCTTCACAGTGGGCTCCCTGG 874
Oy 260 yAla-----GlnGluGluProGlyLeuLeuArgValAlaGlyAspAsnGlyIlePr 277
Db 875 GCGCCCTGGTGGCCACAGAGGGCTGGGCTGCTCGCGCTGGCTGTGACGCGCCATCGC 934
Oy 277 OTPrSerSerAsnAsp---GluLeuPheGlnThrPheCysAspPheProGluIleValAs 296
Db 935 CTGGACCCGAGGAGAAAGAGAGTCTCCAGCCCTTCGACACTTCAGAAATCGTAGA 994
Oy 296 pValSerIleAsnGlnAlaProArgValGlyProAlaGlyLysIleHisArgLeuValThrVa 316
Db 995 CATTTGACATCAAGCAGGCGCCGCGGTTGGCCGCGGAGACAGCGGCTGGTCACTGT 1054
Oy 316 lThrArgMetAspGlyHisIleLeuGluAlaGluPheProGlyLeuProGluAlaLeuSe 336
Db 1055 TACCGAGACAGACAACAGATTTTAAAGCGGAGTCCCAAGGCTGCCAGGGCTCTGTC 1114
Oy 336 rPheValAlaLeuValAspGlyTyrPheArgLeuIleCysAspSerArgHisTyrPheCy 356
Db 1115 GTTGTGGCGCTCGTGGACGGCTACTTCGCGCTGACCAAGGACTCCAGACTTCCTTCTG 1174
Oy 356 sLysGluValAlaProProArgLeuLeuGluGluGluAlaAspValCysHisGlyProI 376
Db 1175 CAAGAGAGGTGGACCGCGAGGCTGTGAGAGAAAGTGGCGAGACAGGCGACGCGCCCAT 1234
Oy 376 eThrLeuAspPheAlaIleHisLysLeuLysAlaAlaGlySerLeuProGlyThrTyrI 396
Db 1235 CACTCTGACTTTCATCAACAAGCTCAAGACTGGGGGCTCAAGCTCCGAGCTCCATGT 1294
Oy 396 eLeuArgArgSerProGlnAspTyrAspSerPheLeuThrAlaCysValGlnThrPr 416
Db 1295 TCTCCGCGCAGCGCCAGAGACTTTGACAGCTTCCTCTCAGTGTGTCAGAAACC 1354
Oy 416 oLeuGlyProAspTyrLysGlyCysLeuIleArgGlnAspProSerGlyAlaPheSerIe 436
Db 1355 CTTTGTGCTGATTAAGAGGCTGCTCATCCGCGCAGCGCCACAGGAACTCTCTT 1414
Oy 436 uValGlyLeuSerGlnProHisArgSerLeuArgGluLeuLeuAlaAlaCysTyrPAsnSe 456
Db 1415 GGTGGGCTCAGCGCAACCCACAGCAAGTCTTCGAGAGCTTCGGCAACCTCTGGGATCG 1474
Oy 456 rGlyLeuArgValAspGlyValAlaAlaLeuThrLeuThrSerCysValaProArgProLy 476

Db 1475 GGGGCTGCACGATAGTGGGGTGGCAGTAGTACCCTACTTCTCGTGTATCCCAAGACCAA 1534
Oy 476 sGlyLysSerAsnLeuIleValValArgArgGlyCysAsnProAlaProAlaProGlyCy 496
Db 1535 AGAAAGTCCAAACCTGATCGGTGCCAGAGAGTCCACAGCCCAACATCATCTTGGT 1594
Oy 496 sSerPro---SerCysCysAlaLeuThrGlnLeuSerPheHisThrIleProThrAspSe 515
Db 1595 TCAGCCCAATCCCAATACAGCTGAGTCAAGATGACATTTTCACAAGATCCCTGACAG 1654
Oy 515 rLeuGluThrPheIleGluAsnLeuGluHisGlySerPheThrLysIlePheArgGlySerAr 535
Db 1655 CTTGAGAGTGGCATGAGAACTGGGCGCATGGGCTCTTACCAACATTTACCGGGGCTGTG 1714
Oy 535 gArgGluValAlaAspGlyGluThrHisAspSerGluValLeuLeuLysValMetAspSe 555
Db 1715 CCATGAGGTGGTGTGATGGGAGGCGCCGAAACAGAGAGTCTCTCAAGGTCATGGATGC 1774
Oy 555 rArgHisArgAsnCysMetGluSerPheLeuGluAlaAlaSerLeuMetSerGlnValSe 575
Db 1775 CAACACAAAGAACTGATGAGATCATTCCTGGAAGCAGCGAGCTTATAGCCAGAGTGC 1834
Oy 575 rTyrProHisLeuValLeuLeuHisGlyValCysMetAlaGlyAspSerIleMetValGl 595
Db 1835 GTACCGGCAATCTGTGCTGCTCCAGCGCGGTGTGCATGCTGGAGACAGCACCATGGTGA 1894
Oy 595 nGluPheValTyrLeuGluValAlaIleAspMetTyrLeuArgLysArgGlyHisLeuValSe 615
Db 1895 GGAATTTGTACACCTGGGGGCCATTACACATATATCTGCCAAACGAGGCGCACCTGGTCC 1954
Oy 615 rAlaSerTyrLysLeuGlnValThrLysGlnLeuAlaTyrAlaLeuAsnTyrLeuGluAs 635
Db 1955 AGCCACTGGAACCTGACAGTGGTGTCAACAGCTGGCTACGCCCTCAACATATCTGGAGA 2014
Oy 635 pLysGlyLeuProHisGlyAsnValSerAlaArgLysValLeuLeuAlaArgGluGly 655
Db 2015 CAAGAGCTGGCCCATGCGAATGTCTGTCCCGGAAGTCTCTGCTGGCGAGGGGCG 2074
Oy 655 yAspGlyAsnProProPheIleLysLeuSerAspProGlyValSerProThrValLeuSe 675
Db 2075 TGAATGGAGCCCGCCCTTATCATACAGTGAAGTCCCTGGGCTCAGCCCGCTGTATAG 2134
Oy 675 rLeuGluMetLeuThrAspArgIleProThrValAlaProGlyLysLeuGlnGluAlaGl 695
Db 2135 CCTGGAGATGCTCACAGCAGAGATCCCTGGGTGGCCCGCAGTGTCTCCGGAGGCGCA 2194
Oy 695 oThrLeuCysLeuGluAlaAspLysTyrGlyPheGlyValaThrThrTrpGluValPheGl 715
Db 2195 GACACTTGAAGTGGAAAGCTGACAGTGGGCTTCGGCGCCACGGTCTGGGAAGTGTAG 2254
Oy 715 nArgGlyProAlaHisIleThrSerLeuGluProAlaLysLysLeuLysPheTyrGluAs 735
Db 2255 TGGCCTCAACATGCCCATCACTGCTGGATCTGTGTAAGAACTCCAAATTTATAGGA 2314
Oy 735 pGlnGlyLysLeuProAlaLeuLysTyrTrpGlnLeuAlaGlyLeuLeuThrGlnCysMe 755
Db 2315 CCGGAGACAGCTGCGGCGCCCAAGTGAACAGACTGGCCCTGTGATTCACAGAGTGCAT 2374
Oy 755 tAlaTyrAspProGlyArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLe 775
Db 2375 GGCCTATGAGCGGCTCCAGAGGCTCTCTTCCAGAGCGGTATTCGTGACTCAATAGACT 2434
Oy 775 uIleThrSerAspTyrGlyLeuLeuSerAspProThrProGlyIleProSerProArgAs 795
Db 2435 CATCTCTTCAAGATATAGTCTCTTCAGACCCACACCTGGTGGCCCTGGCACCTGTGA 2494
Oy 795 pGluLeuCysValAlaGlyValGlnLeuTyrAlaCysGlnAspProAlaIlePheGluGl 815
Db 2495 TGGGCTGTGG---AATGGTCCCAAGCTTAATGCCCTCAAGAACCCACAGATCTTGAGGA 2551
Oy 815 uArgHisLeuLysTyrIleSerLeuLeuGlyLysGlyAsnPheGlySerValGluLeuCy 835
Db 2552 GAGACACTCAAGTATCATCTCACAGCTGGGCAAGGCAACTTTGGCAGCGTGGAGCTGTG 2611

```

OY 835 sarGTYrAsPProLeuGlyAspAsnTrpGlyProLeuValAlaValLysInLeuGlnH 855
    |||||
Db 2612 CCGTATAGACCCGCTAGCGCACATACAGGTGCCCGGTGGCCGTGAACAGCTGCAGCA 2671
    |||||
OY 855 sSerValProAspGlnGlnArgAspPheGlnArgGluGlnGlnLeuLysAlaLeuH 875
    |||||
Db 2672 CAGCGGCGCACAGCAGAGAGGACTTTCACGGGAGATTCAGATCTCTCAACACAGTGA 2731
    |||||
OY 875 sSerAspPheLeuValLysTrpArgGlyValSerTrpGlyProGlyArgGlnSerLeuAr 895
    |||||
Db 2732 CAGTATTTCATTGCAAGTATCTGGTGTCAGCTATGCGCGGCGCACAGACCTGCGC 2791
    |||||
OY 895 GLeuValMetGluTrpTrpLeuProSerGlyCysLeuArgAspLeuGlnArgHisArgG 915
    |||||
Db 2792 GCTGGTCAATGAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2851
    |||||
OY 915 Y---LeuHisThrAspArgLeuLeuLeuLeuPheAlaTrpGlnIleCysLysGlyMetGlu 934
    |||||
Db 2852 GCGGCTCGATGCCAGCCGCTCTCTCTATCTCTGCGCAGATCTGCAAGGGCGATGAGTA 2911
    |||||
OY 934 rLeuGlyAlaArgArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGluSe 954
    |||||
Db 2912 CCGTGGCTCCCGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2971
    |||||
OY 954 rGluAlaHisValLysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAs 974
    |||||
Db 2972 CGAGGCACACGTCAAGATGCTGACTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3031
    |||||
OY 974 pPYrTrpValValArgGluProGlyValInSerProIlePheTrpTrpAlaProGluSerLe 994
    |||||
Db 3032 CTACTAGCGTGGTCCGCGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 3091
    |||||
OY 994 uSerAspAsnIlePheSerArgGlnSerAspValTrpSerPheGlyValValLeuTrpG 1014
    |||||
Db 3092 CTGCGAACACATCTCTCTGCGCAGTCAGACGTCGAGCTGCGGCTGCTGCTGCTGCTGCTG 3151
    |||||
OY 1014 uLeuPheTrpTrpCysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetG 1034
    |||||
Db 3152 GCTCTTCACTACTGCGCAAAAGCTGACGCCCTCGCGGAGTTCTCTCGGATGATGAGG 3211
    |||||
OY 1034 YProGluArgGluGlyProProLeuGlyCysArgLeuGluLeuAlaGluGlyArgAr 1054
    |||||
Db 3212 ATGAGACGGGAGTGTCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3271
    |||||
OY 1054 GLeuProProProProTrpTrpArgValGlnGluLeuMetGlnLeuGlyStrpAl 1074
    |||||
Db 3272 GCTGCGCGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3305
    |||||
OY 1074 aProGluProHisAspArgProAlaPheAlaTrpLeuSerProGluLeuAspProLeuTr 1094
    |||||
Db 3306 -----ACGCCCGGAGGCGTACGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3346
    |||||
OY 1094 pArgGlyArgProGly 1099
    |||||
Db 3347 GCGCGGCGTCTGCGGC 3362
    |||||

```

RESULT 8 AC073700 218208 bp DNA linear HTG 29-JUN-2000
 AC073700 LOCUS Mus musculus clone RP23-162012, WORKING DRAFT SEQUENCE, 16
 DEFINITION Mus musculus clone RP23-162012, WORKING DRAFT SEQUENCE, 16
 ACCESSION AC073700
 VERSION AC073700.1 GI:8810317
 KEYWORDS HTG: HTGS_PHASE1; HTGS_DRAFT.
 SOURCE Mus musculus.
 ORGANISM Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 218208)
 AUTHORS DOE Joint Genome Institute.
 TITLE Sequencing of Mouse
 JOURNAL Unpublished

REFERENCE

2 (bases 1 to 218208)
 DOE Joint Genome Institute.
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 Center: Joint Genome Institute
 Center Code: JGI
 Web site: <http://www.jgi.doe.gov>

Project Information
 Project Name: 1799428
 Center clone name: RPCI-23_162012

Summary Statistics

Consensus quality: 204328 bases at least Q40
 Consensus quality: 210800 bases at least Q30
 Consensus quality: 212488 bases at least Q20
 Estimated insert size: 22200; agarose-1p estimation
 Estimated insert size: 216708; sum-of-contigs estimation
 Quality coverage: 9.37 in Q20 bases; agarose-1p estimation
 Quality coverage: 9.6 in Q20 bases; sum-of-contigs estimation.
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 16 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 1092: contig of 1092 bp in length
 1 1192: gap of unknown length
 1 1193: contig of 1373 bp in length
 1 2565: gap of unknown length
 1 2566: gap of unknown length
 1 3711: contig of 1046 bp in length
 1 3811: gap of unknown length
 1 3812: contig of 1744 bp in length
 1 5555: gap of unknown length
 1 5556: gap of unknown length
 1 5656: gap of unknown length
 1 9015: contig of 3359 bp in length
 1 9115: gap of unknown length
 1 12524: contig of 3410 bp in length
 1 12525: gap of unknown length
 1 12624: gap of 5691 bp in length
 1 12625: contig of 5691 bp in length
 1 18316: gap of unknown length
 1 18316: gap of unknown length
 1 18416: contig of 6884 bp in length
 1 25300: gap of unknown length
 1 25300: gap of unknown length
 1 25400: contig of 9459 bp in length
 1 34859: gap of unknown length
 1 34859: gap of unknown length
 1 48111: contig of 13153 bp in length
 1 48111: gap of unknown length
 1 48212: contig of 11464 bp in length
 1 59675: gap of unknown length
 1 59676: gap of unknown length
 1 59775: gap of unknown length
 1 79239: contig of 19464 bp in length
 1 79339: gap of unknown length
 1 79240: gap of unknown length
 1 79340: contig of 23420 bp in length
 1 102759: gap of unknown length
 1 102859: gap of unknown length
 1 102860: contig of 27114 bp in length
 1 129974: gap of unknown length
 1 129974: gap of unknown length
 1 130073: gap of unknown length
 1 130074: contig of 34115 bp in length
 1 164189: gap of unknown length
 1 164289: contig of 53920 bp in length.

FEATURES

source

1. 218208

/organism="Mus musculus"

/db_xref="taxon:10090"

/clone="RP23-162012"

/clone_lib="RPCI mouse BAC library 23"

51066 a 56856 c 56653 g 52132 t 1501 others

BASE COUNT

ORIGIN

Alignment Scores:

Pred. No.: 5,71e-240 Length: 218208
 Score: 4091.50 Matches: 1072

Percent Similarity: 32.75%
 Best Local Similarity: 32.69%
 Query Match: 69.82%
 DB: 2
 Conservative: 2
 Mismatches: 24
 Indels: 2188
 Gaps: 24

US-09-397-967-16 (1-1099) x AC073700 (1-218208)

QY 1 MetAlaProSerGluGluThrProLeuIleProGlnArgSerCysSerLeuSer 20
 Db 110153 ATGGACCTCCAACTGAGAGACACCTCTGATCCCTAGCGCTCTTGACCTCTCATCC 110212
 QY 21 SerIuAlaGlyAlaLeuHisValLeuLeuProArgIleProGlyProGlnArg 40
 Db 110213 TCACAGGCGAGAGCCCTGCATGTGCTCTCTCCCGGGAGCTGGCCCTCCACCGCA 110272
 QY 41 LeuSerPheSerPheGlyAspIleuValGluAspLeuValArgAlaAlaVal 60
 Db 110273 TTGTCATTCCTCTTTGGGGACTACTGCTGAGGATTTATGTGGAGCTGCCAAGGCC 110332
 QY 61 CysGlyIle----- 63
 Db 110333 TGTGTGA-GTGGTCCCTAGATGACCATGCTATATAGCAACATCTATTGAGAAG 110391
 QY 63 ----- 63
 Db 110392 GGATGCTGGGCTTTTGAAGGGCTGTGAATTTGGTAAGGAGATCATGAGATCAGCCG 110451
 QY 64 -----LeuProValIleHis 68
 Db 110452 TGTAGTTGACTTGGTGAAGTATGTCCTTCAAGCATCTGCTGTTATCAT 110511
 QY 69 SerLeuPheAlaLeuAlaThrGluAspPheSerCysTrpPheProSerHisIlePhe 88
 Db 110512 TCCTTTTGGCTGGCCACTGAGGACTTCTGTGCTGTTTCCCAAGCACATCTTC 110571
 QY 89 CysIleGluAspValAspThrGlnValLeuValIleArgLeuArg----- 103
 Db 110572 TGCATGAGGAGGTGAGACCTCAAGCTTGTGCTACAGGCTTACGAGAGAGAGTGAACCC 110631
 QY 103 ----- 103
 Db 110632 TTGATTTATCTTCAAGGATGATTTTCTATCTATAGTATAGAGGCTTGAACAG 110691
 QY 103 ----- 103
 Db 110692 GAGGTTGGGCTGTTGCTGTAGGACACACTTCACTTCACTGACCTGGG 110751
 QY 104 -----PheTrpPheProAspTrp 109
 Db 110752 AAGCCAGAGCTGAGTGGCCACCCCATTTCTGAAACCCACAGCTTTTATTTCCCTGAC 110811
 QY 110 PheGlyLeuGluThrCysHisArgPheGlyLeuArgIleAspLeuThrSerAlaIleLeu 129
 Db 110812 TTTGGGCTGGACATGTGACCGCTTTGGCTGGCAAGATTTGACACAGTGCATCTT 110871
 QY 130 AspLeuHisValLeuGluHisIlePheAla----- 139
 Db 110872 GACTTACACGTTTTAGAACATCTTTGCTCAGGTGGTATTCTTCCGTGGGGAAGCTC 110931
 QY 139 ----- 139
 Db 110932 AAGCCAGGATGGGGAGCCCAAGGTAAACAAATCCAGACCCAGAGCGAAGCTCACT 110991
 QY 140 -----GlnHisArgSerAspLeuValSerIleArgLeuProVal 152
 Db 110992 AACCACTTCTGTGGCTGACACCAAGACACAGTACCTGTGGTGGCGCTCCCGGT 111051
 QY 152 IGlyLeuSerMetIleGluGluIlePheLeuSerLeuAlaValLeuAspLeuAlaI 172
 Db 111052 GGGCTTAGCTGAAAGAGCGAGGAGATTCCTGAGCCCTGCTGACACTTGGCCCA 111111
 QY 172 mMetAlaArgGluGlnIleGlnArgProGlyGluLeuLeuIleThrVal----- 188

Db 111112 GATGCTCTGTGACAGAGCCAGCCAGAGAGAGCTGTGAGAGGTCAGGTGAGGCC 111171
 QY 188 ----- 188
 Db 111172 TCGGGCCACCCCTTGTGTCTTCTTTGGCTGGAGTCAAGGCCATTGTTTCCCT 111231
 QY 188 ----- 188
 Db 111232 ACTGAAACATACCTCTTCCCACTTGTGAGGACACCTCTCGCTGAGCTGCCCTG 111291
 QY 188 ----- 188
 Db 111292 CCGAAGCTCCCTATGTCTGTCTGTCTTCCGGGCCCCCTCACCACTTAAGAC 111351
 QY 189 -----SerTrpIleValGlyLeuProSerLeuArgAspVal 202
 Db 111352 TCGGGCTGCCCCCCCCCCACAGTTACAAAGCTGTCTGCGCCAGCCTGGCGATGTGA 111411
 QY 202 IeGluGlyGlnAspPheValThrArgArgIleArgArgThrValIleLeuAlaLeu 222
 Db 111412 TCCAGGGCCAGAACTTCTGTACAGCAGGCGCATCCGAGAGCGGTGTGGCGCTGC 111471
 QY 222 euProCysGlyArgLeuProGlyArgPro-TyrAlaLeuMetAlaIleLeuAsp 241
 Db 111472 G-CCCTGTGTGCTGCTGCGCAGGCGGACCGCTACGCGCTCAAGGCCAAGTATTTTGAC 111530
 QY 242 LeuGluArgLeuHisProAlaAlaThrThrGluThrPheArgValIleLeuProGlyVal 261
 Db 111531 CTGGAGCGGCTAACATCCAGCGGCGCACACAGAGACTTCCGTGTGGGGCTCCCGGGCGCC 111590
 QY 262 GlnGluGluProGlyLeuLeuArgValAlaGlyAspAsnGlyIleProTrpSerSerAsn 281
 Db 111591 CAGAGAGGCGGGGCTTCTGTGTGTGGGGGAGAACGCACTTCTGAGCTCCGG 111650
 QY 282 AspIle----- 283
 Db 111651 GACCA-GGAGGTGCTGGGGCTGGGCTTGAAGAGCTGGTCCCGGGAGGCGTGGCGGT 111709
 QY 283 ----- 283
 Db 111710 TAGTAAAGAACCCCTCCGAACATCAGAAAGTGGCTAGTCCGGCGGAGTGTGCA 111769
 QY 283 -----SerTrpIleValGlyLeuProSerLeuArgVal 261
 Db 111770 CTAGGGGCGAGGGGGGGGCTCAATTAAAGTGGGTCCGGGATGACCTACCGAGGA 111829
 QY 283 ----- 283
 Db 111830 CTGAGCTGGGGCGGTGACTAGAAAGGTGAGACTAGATGAGGGGGCGGAGCGAGAG 111889
 QY 283 ----- 283
 Db 111890 GCGATAGTATGATTTGATGATCGGTGGGTGTGACCCGACAGGCTGTACGTCCCTAAC 111949
 QY 283 ----- 283
 Db 111950 CTGTCGCCCGACCCATCCACCCCTCCCGCCGTAGTGGAGGTGGGAAGAGAG 112009
 QY 283 ----- 283
 Db 112010 AAGGAGAGCTGTGTGGTCAAGGTCTGTAATCTGTGCTTCTGCTGCTCTTCT 112069
 QY 283 ----- 283
 Db 112070 CTCTTTTATATGTTTTTGGAGAGAGGTTTTCTGTATAGCTTGGCTGTCTGGA 112129
 QY 283 ----- 283
 Db 112130 ACTCACTTTGTAGACAGGCTGGCTGGAAGTCAAGAAATCCGCTGCTGTCTGCA 112189
 QY 283 ----- 283
 Db 112190 GTGCTGTGATTAAGCGGTGGCGCACTACACCAAGCTACTGCTGTCTTCTCTCAG 112249

Oy 284 LeuPheGlnThrPheCysAspPheProGlnIleValAspValSerIleAsnGlnAlaPro 303
 Db 112250 CTTTTCAGACCTTCTGTGACTTCCGAAATCGTGGATGTCAGTCATCAAGAGGCCCA 112309
 Oy 304 ArgValGlyProAlaGlyIleuHisArgLeuValThrValThrArgMetAspGlyHisIle 323
 Db 112310 CGTGTGGCTCCGCGAGGGAACACCGGCTGTACCGTCACCGAGATGACGGCCACATC 112369
 Oy 324 Leu----- 324
 Db 112370 CTGGTGGCTCTGGACCCGTTATTCTGTGCTGATCCTGGAAAGGGGTCCGTT 112429
 Oy 325 -----GlnAlaGluPheProGlyLeuProGluAlaLeu 336
 Db 112430 GTTCTGAGAGTCTCTCGCGCAGGAAGCGAGTTTCCGGGGCTGCTGAGCGCTGTC 112489
 Oy 336 rPheValAlaLeuValAspGlyTyrrPheArgLeuIleCysAspSerArgHisTyrrPheCys 356
 Db 112490 CTTCGTGCCCCCTGTGTGATGGGTACTTCCGCTGATCTGGACATCCAGGCATTATTCTC 112549
 Oy 356 sIysGluValAlaProProAlaGlyLeuGluGluAlaAspValCysHisGlyProIle 376
 Db 112550 CAAAGAGGTGGCGCGCGCGCGCTGTGAGAGAGAGCGGAGCTGTGCATGAGCCAT 112609
 Oy 376 eThr----- 377
 Db 112610 CACGTTAGAGCTAGTGGTGGGTACAGTGGGAGGGGCTGGCAGCAGAGCATGCTCG 112669
 Oy 378 -----LeuAspPheAlaIleHisIysLeuValAla 388
 Db 112670 GTTGGACCTGGCTGCTGCTTGGCTTGAAGTTAGACTTGGCATCCAAAGCTGAAGCGCG 112729
 Oy 388 IagIySerLeuProGlyThrTyrrIleLeuArgArgSerProGlnAspTyrrAspPheIle 408
 Db 112730 CTGGCTCCCTCCAGCAGCATTACTTCCGCCGAGCCCGAGAGATGATACAGCTTTC 112789
 Oy 408 euleuThrAlaCysVal----- 413
 Db 112790 TTCTTACCGCTGCGCT-CCAGGTTGTTTACTGCTGTGTCAGAGATGGCTTCTGAGCA 112848
 Oy 413 ----- 413
 Db 112849 AGTGACACTAGTGGCAGCAGAAAGCGCTTCTATTCTTGAAGACACTGGCC 112908
 Oy 413 ----- 413
 Db 112909 TGTAACTCAGATTAAAGAGACTGAGCAGATCAAGTTTACAGGCTGCTGAGCTAT 112968
 Oy 413 ----- 413
 Db 112969 ATAGAGAAATTATGTTTCTCATGAGGAAAGGTTTTTCCACCAAGCTGAGGGTCT 113028
 Oy 413 ----- 413
 Db 113029 CTTTATCATGATTCACAGCATGAGAGACAGACATGACACCATTAATCGGCTCG 113088
 Oy 413 ----- 413
 Db 113089 ACCTCCACAGTGCAGGTGATCTGATCTACACAGAAACACATTAAGTAATGATTT 113148
 Oy 413 ----- 413
 Db 113149 TAAGAACTCGAAGAGCTCAGAAACAGTCTCAGCATGGCGGGGGGGGAGAG 113208
 Oy 413 ----- 413
 Db 113209 CATCTTAGAGTTTGAGACAGAACTGAGCATCACTTCCAGAAAGAGAGGGGT 113268
 Oy 414 -----GlnThrProLeuGlyP 419
 Db 113269 GGGTAGCTGTGCTCCACATCTGGCTCAGTGCAGGTCTCTTTCAGACACTCTCTTGCGCC 113328

Oy 419 roAspTyrrLysGlyCysLeuIleArgGlnAspProSerGlyAlaPheSerLeuValGlyL 439
 Db 113329 CCGACTACAGAGGCTGCTCATCCGCCAGAACCCAGGGGGGCTTCTCCCTGGTGGCC 113388
 Oy 439 euSerGlnProHisArgSerLeuArgGluLeuAlaIleCysTrpAsnSerGlyLeu 459
 Db 113389 TCAGGACGCCCAACAGAACCTGCGGAGCTGTTCAGCTGTGGAATCTGGGCTGC 113448
 Oy 459 rValAspGlyAlaLeuTyrrLeuThrSerCysAlaProArgProIlys----- 476
 Db 113449 GAGTAGAGGTCTGCTCCCTGAACCTAATCATCTGCTGGCTCCAGACCAAGGGTAGT 113508
 Oy 476 ----- 476
 Db 113509 CAGTTCTCCCTAACACACCGCGAGCTTGGGTGTAGCAGAAACCAAGACCATGG 113568
 Oy 476 ----- 476
 Db 113569 CGATTAGACAGAGAACACGACTTTTGGCACAGTCTTCAGACATCAAGTCAATTATT 113628
 Oy 476 ----- 476
 Db 113629 CAGTTGGCGCGCTTCTTCCATGGGTGATGCCACTAAGGCGAACCCCTGCTGGACA 113688
 Oy 476 ----- 476
 Db 113689 CATCTCCCTCTGTGTGTCTGTGCTTAACAAAGACCAGCGCGCTGAGGAGAACCCAGT 113748
 Oy 476 ----- 476
 Db 113749 ACGTTCAATGTTCAANTCAGGCCCAAGTCAAGATCTCAGAGCGACTTCCGGGGCGCC 113808
 Oy 476 ----- 476
 Db 113809 CCAAGCTAAGTAGAGAGTGAATCTTGAACCTTGTGCTGACTGCTTCTT 113868
 Oy 477 -----GlyIysSerAsnLeuIleValAlaArgArgIysAsnProAlaProAlaP 494
 Db 113869 GCCATCCAGAAAGTCCAAATTTATGTGTGGTGGAAAGGGGTCTCACCCGGCGCTGCC 113928
 Oy 494 roGlyCysSerProSerCysAlaLeuThrGlnLeuSerPheHisThrIleProThr 514
 Db 113929 CTGGCTGCTCCCGCTGCTGCTGTGCGCTGACAGCTGAGCTTCCACAAATTCACAGG 113988
 Oy 514 sPserLeuGluTrp----- 518
 Db 113989 ACAGCTGAGTGGGTAGGGTCCCAAGGTGGAAGCCAGGGCGGCTTTTGGAGG 114048
 Oy 519 -----HisGluAsnLeu 522
 Db 114049 ACGTGAAGCCACTTCAACTGTCCTGCGCTGGGTGTCTCCCGCAGCACGAGAACCTG 114108
 Oy 523 GlyHisGlySerPheThrIysIlePheArgIysArgArgGluValAlaAspGlyGlu 542
 Db 114109 GGTACGGTCTTTTACCAATCTTCCGTGGCGGAGGGGAGGTGCTGATGTGTGAG 114168
 Oy 543 ThrHisAspSerGluValLeuLeuLysValMetAspSerArgHisArgAsnCysMetGlu 562
 Db 114169 ACACATGACTGGAAGTCTCTCTGAGAGTATGATGACATCCGAACTCGAATCTGATGAG 114228
 Oy 562 ----- 562
 Db 114229 GTGCGAGGGAGCAGAGAGGCTTGCAGAACTGCATGGAATCTGGACCTGGCAGAGG 114288
 Oy 562 ----- 562
 Db 114289 GTTACTGGAACCTAGAACTTCAAGACACAGACTGGGGGAGCATAGTGGGACTCCATCAACC 114348
 Oy 562 ----- 562
 Db 114349 AAGTAGATTAGCATCTGTGATTGTGTAAACACAGCATGAGAGGCTGGGAGG 114408
 Oy 562 ----- 562

QY 778 ----- 778
Db 116627 CATCTCAGGCTGATTTCTCCCTGCTTCCACCATCATGGGTGCGGATTTGGTTTATTT 116686
QY 778 ----- 778
Db 116687 TGTATTCTTAATTTTGTGACACAGATGTTCTCATCTGAGGGATTCAGATCAGACT 116746
QY 778 ----- 778
Db 116747 GGCCTCAGTGCATGGCTGTCTCTGCGGTCCTTTTGTCTGAGCCGTTGGGTCACA 116806
QY 778 ----- 778
Db 116807 CAGAGTGTCTTTCCCGGACAGTGTGCTTCTCATTTCCCATTAAGTACTGTGGCG 116866
QY 778 ----- 778
Db 116867 TCTCATTAACGACAGAGACAGAGTTTCAGGTGAGTTTACACCATTCAGTTGGG 116926
QY 779 ----- 779
Db 116927 ATTGAGGACGAGACTTACCATTTCCCATCTTCTCAGGTATCCCGCTTACATTTACG 116986
QY 781 ----- 781
Db 116987 AGCTCTCTCAGACCCGACACCTGGCATCCGAGTCTCTGAGATGAGCTGTGC ---GCTG 117043
QY 801 ----- 801
Db 117044 GCGCCAGGCTCTATGCTGCTGACAGACCCGCGCATATTGAGAGAGACACCTTAAGTACA 117103
QY 821 ----- 821
Db 117104 TCTCTTTGCTGGGCAAGGTAGTGGCGGCGCATGTGGGGAGGAACGTGGTGGTGGAT 117163
QY 827 ----- 827
Db 117164 GGGTCAGGTGAGACATGCTCTCATCTTCATCCACAGGGCAACTTTGGACGGAGCTG 117223
QY 835 ----- 835
Db 117224 TGCCGCTATGACCCCTGGGGGACATACGAGACCCCTGTTGGCAGTGAACAGCTACAG 117283
QY 855 ----- 855
Db 117284 CACAGCGGGGACAGACAGACAGAGGAGCTTCAGCGGAGATTGATCTTAAGGCTCTG 117343
QY 875 ----- 875
Db 117344 CACAGCGGACTTCATCTCATGATGACGAGGAGTCCAGGAGATTGATCTTAAGGCTCTG 117403
QY 889 ----- 889
Db 117404 CATCTCGGGAAGGGTGTGAGTCCGCTTCTACCTTTCCAGTAGGGCCCTGTTGACA 117463
QY 889 ----- 889
Db 117464 AGGTCTTAATCTCCATGATCCAGCTTCTATAGCTGAGGTTGAACGATCATACAC 117523
QY 890 ----- 890
Db 117524 CCAGTAGAAGGTCTCGAAGCTACTGTGCTTCGCGAGTCCGCAAGGCTCGGCTG 117583
QY 897 ----- 897
Db 117584 GTATGAGTACCTGCCAGCGGCTCTCTCGAGACTTCTTCAGGCGCATTCGCGCGG 117643
QY 916 ----- 916
Db 117644 CTGACACGACGCGCTACTGCTGTCTGCTGAGATCTGCAAGGTTGGAGACATCGCGG 117703
QY 930 ----- 930
Db 117704 AGATCGGATGGGCTCTGCGGTCGCGGATCCCGGCGCCAGCTGATCATGCTCTCT 117763
QY 931 ----- 931
Db 117764 GCAGGCGATGAGTACTGGGTGCGCGGCTGCTGACACCTGATCTGCTGCGCGCA 117823
QY 949 ----- 949
Db 117824 CATCTGTGTGAGAGCGAGCTCATGTGAAGATGCGGAGCTTGGCTGCTGATAGCTGCT 117883
QY 969 ----- 969
Db 117884 GCGCTGGGAAGACTACTGATGCTGCTGCGGAGCTGCGCAAGGCCCATCTTTTGGTA 117943
QY 988 ----- 988
Db 117944 CAAGACCAAGACTCCAGAGTACAGCCCACTGAGCCCACTCCGCACTACAGCTT 118003
QY 988 ----- 988
Db 118004 AAACGCTCAGACCCAGCTCCACCTCAGTAGCTTGGCTCGCAAGGCTTGGCCACA 118063
QY 988 ----- 988
Db 118064 GCAGGCTCAGCCCTATAGCCAGATCCAGAACCCATGCTCAACACTTACAGATACC 118123
QY 988 ----- 988
Db 118124 CCTCTCACTTACTGATGACGACTACCCACCACCCACAGAGTATGCTCAGCTGTTTA 118183
QY 988 ----- 988
Db 118184 GCGACGCCCCCATGTTGTGCTGCTCATCTCTCTTGACAGTCTGCAAGGCCGTCACAG 118243
QY 988 ----- 988
Db 118244 GTCTCTTTTCTAGCTCCGCTCCCTCGGAGGCGCTGCTTACAGCTATAGGGCTCCG 118303
QY 988 ----- 988
Db 118304 CCGCTTAGCGCTCTTACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 118363
QY 988 ----- 988
Db 118364 CCGCTTAGCATTTAGCGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 118423
QY 988 ----- 988
Db 118424 CTGACCCACATGAAGCCCGCTGCTTAACTTGAAGTCCCTGCTGCTGCTGCTGCTGCT 118483
QY 988 ----- 988
Db 118484 CACAGGTACACAGACGCTGTGAGTGGCTTGTGGGGCGCTGCTGAGGAATCTCC 118543
QY 988 ----- 988
Db 118544 TGTTCAGCCCTTCTCTACCCAGAGCTGSCCATTTACATTTAGAGACTCCCTCC 118603
QY 988 ----- 988
Db 118604 AGGCGCATGTTTACACCTGCTCCACTCTCTCATCTGCGCACACACACACTCTGTCCGC 118663
QY 989 ----- 989
Db 118664 TCTCAGTATGCGCCGAGGCCCTATCTGACACATCTTCTCCGCAATCTGAGCTGT 118723
QY 1006 ----- 1006
Db 118724 GAGCTTGGAGTGTGTGTGAGAGCTTTCACCTTACCTGCGACAGAGCTGAGGCCATTC 118783
QY 1026 ----- 1026
Db 118784 TCTGACACGACGCGCTACTGCTGTCTGCTGAGATCTGCAAGGTTGGAGACATCGCGG 117703

Db	119863	GCTATGCAAGCTGTGCTGGGGGGCCACCCAGCCGACAGACCCGCCACGCTTGCGCACACCTTGAG	119922
Oy	1087	rProglInleuaspProleutrParaglYarProglY	1099
Db	119923	CCCCAGCTGTGACGCGCTGTGGCGCTGTGAGAACCCCGA	119959
RESULT 9			
AC073750			
LOCUS			
DEFINITION	AC073750	223734 bp	DNA linear HTG 29-JUN-2000
	Mus musculus	clone R123-330D8,	WORKING DRAFT SEQUENCE, 28 unordered
ACCESSION	pieces.		
VERSION	AC073750		
KEYWORDS	AC073750.1	GI:8810367	
SOURCE	HTG: HTGS_PHASE1; HTGS_DRAFT.		
ORGANISM	Mus musculus.		
	Mus musculus		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus		
	1 (bases 1 to 223734)		
	DOE Joint Genome Institute.		
	Sequencing of Mouse		
	Unpublished		
	2 (bases 1 to 223734)		
	DOE Joint Genome Institute.		
	Direct Submission		
	Submitted (29-JUN-2000)		
	Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
	-----Genome Center		
COMMENT			

```

Project Information
Center Project Name: 1863672
Center clone name: RPCR-23_330D8
-----
Summary Statistics
Consensus quality: 196566 bases at least Q40
Consensus quality: 208241 bases at least Q30
Consensus quality: 210965 bases at least Q20
Estimated insert size: 206800; agarose-fp estimation
Estimated insert size: 221034; sum-of-contigs estimation
Quality coverage: 8.01 in Q20 bases; agarose-fp estimation
Quality coverage: 7.54 in Q20 bases; sum-of-contigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 28 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1
1026      1027: contig of 1027 bp in length
1128      1127: gap of unknown length
2162      2161: contig of 1034 bp in length
2262      2261: gap of unknown length
3710      3709: contig of 1448 bp in length
3810      3809: gap of unknown length
4864      4864: contig of 1055 bp in length
4865      4864: gap of unknown length
4965      4965: contig of 1021 bp in length
5986      6085: gap of unknown length
6086      7691: contig of 1606 bp in length
7692      7791: gap of unknown length
7792      10317: contig of 2526 bp in length
10318     10417: gap of unknown length
10418     11756: contig of 1339 bp in length
11757     11856: gap of unknown length
11857     13223: contig of 1367 bp in length
13224     13323: gap of unknown length
13324     16114: contig of 2791 bp in length
16115     16214: gap of unknown length
16215     20778: contig of 4564 bp in length

```

Db 124399, CGATGCTGGGTTTTCAGCGGCTGGCAATTTGGTAAGGAGAGATCATGGGATCAGCCG 12445

Qy	64	-----	LeuProVal ¹ Trp ¹ His	68
Db	124459	TGTAGTTGACCTGGGTGAACTGATGTGCCCCATCCCTTCAGAGCATCCCTGGTTATCAT	124518	
Qy	69	SerLeuPheAlaLeuAlaIhrGluAspPheSerCysTrpPheProPheSerHisIlePhe	88	
Db	124519	TCGGCTTTCGGCTCGGCCACTGAGGACTTCCTCTGTGTGGTTCCGCCAAGGCACATCTTC	124578	
Qy	89	CysIleGluAspValAspTrpGluValLeuVal ¹ TrpArg	103	
Db	124579	TGCATAGAGGACGCGTGCACACTCAAGTCCTGTGTCTACAGGCTCAGGTAGAGAGTGAACCCC	124638	
Qy	103	-----	103	
Db	124639	TTGATTATCTTCAGGGTAGTATTTCATTACACTGATATGTAGAGGCTTTGAACACAGA	124698	
Qy	103	-----	103	
Db	124699	GAGCTTGGGCTGTGTGGCTAGTGGCCACACTGCACCTTCAGTGAACCCAACTTGGGCTCGG	124758	
Qy	104	-----PheTrpPheProAspTrp	109	
Db	124759	AAGCCAGAGCTGAGTGGGCCACCCCATTTCTGAAGACCACAGCTTTATTTCCCTGACTGG	124818	
Qy	110	PheGluLeuGluTrpCysHisArgPheGluLeuArgLysAspLeuHisSerAlaIleLeu	129	
Db	124819	TTTGGCGTGGAGACATGTCAACGCTTTGGGCTGGCGAAAGATTGACAGTGCATCTTT	124878	
Qy	130	AspLeuHisValLeuGluHisIleuPheAla	139	
Db	124879	GACTTACACGTTTTAGAACATCTCTTGGCTCAGGTGTGTATTCTTCCGGTGGGAAGGCTC	124938	
Qy	139	-----	139	
Db	124939	AACCCAGGGATGGGGAGCCCAAGGTAAACAATCCAGCACCCAGAGACGGAACCTCAGCT	124998	
Qy	140	-----GlnHisArgSerAspLeuValSerCyl ¹ ArgLeuProVal	152	
Db	124999	AACCACTCTTCTGTGGCTGACCCAGCCAGCAGTGAAGCTGTGAGTGGGCGCTCCGGGT	125058	
Qy	152	IleGluSerMetLysGluGlnGluPheLeuSerLeuAlaValLeuAspLeuAlaGlu	172	
Db	125059	GGGCTTACATGAAGAGACAGGAGGAGGAGTTCCTGAGCTGGCCGTGGACTTGGCCCA	125118	
Qy	172	metAlaIar ¹ gGluGlnAlaGlnArgProGluLeuLeuLysThrVal	188	
Db	125119	GATGGCTCGTAGAGCAGGCCAGCGCCAGGAGAGCTGCTGAAGAGCGTCAAGTGAAGGGCC	125178	
Qy	188	-----	188	
Db	125179	TGCGGCGACCCCTTCGTGCTCTTCTTTGGCTGCGGATCTAGGCGCTATCGTTTCCTTT	125238	
Qy	188	-----	188	
Db	125239	ACTGGAACAATACCTCCCTCCCACTGTGTGACGCCACCCCTCTCGTCAGGCTGCCCTG	125298	
Qy	188	-----	188	
Db	125299	CCGAAGAGTCCCTATGTCTGTCTGTGTGCTTCGCGGCCCTCAGCACACCTTAAGAGC	125358	
Qy	189	-----SerTrpLysAlaCysLeuProPheSerLeuArgAspVal	202	
Db	125359	TGCGCGTGGCCCCCCCCACAGATTAAAGAGCTGTCTGCGGCCAGCGCTGGCGCATGGGA	125418	
Qy	202	LeuGlnGluGlnAsnPheValThrArgArgArgIleArgArgThrValValLeuAlaLeuL	222	
Db	125419	TCCAGGGCCGAACCTTGATGACACGGAGCGCATCCGACGAGACCGTGGTCTTTGGCGGTGC	125479	
Qy	222	eupProCysGlyArgLeuProGluArgPro ¹ TyrAlaLeuMetAlaLysTrpIleLeuAsp	241	
Db	125479	G-CCGTGTGGTGGCTGCCAGCGCCAGCCGCTACGGCTCATGTGGCCAAAGTATATTTGGAC	125539	
Qy	242	LeuGluArgLeuHisIleProAlaIleThrThrGlnGluThrPheArgValGluLeuProGluVala	261	

[illegible]

Db 126617	CACGTTAGGAGCTAAGTTGGGGTCCACAGTGGGAGTGGGGCTGGGACGCGAGGACGTGTGG	126617
OY 378	-----LeuAspPheAla11GhiIAsylLeuAla388	
Db 126677	GTTGGACCTGGCTGTGTCTGTGGCCAGGTTAGACTTTGGCATCCACAGCTGAAGGCCG	126736
OY 388	IaGlySerLeuProGlyThrIrrYrIleLeuArgArgSerProGlnAspYrAspSerPheI408	
Db 126737	CTGGCTCCCTCCGACGAGGACCTACATTCTCCGCGGACGCCGACGAGACTATGACAGCTTTC	126796
OY 408	euleuthrAlaCysVal-----413	
Db 126797	TTCTTACCGGCTCCGT-CCAGGTTGGTTTACTGCTGGTCCAGAGATGGCTTCTCGAGGA	126855
OY 413	-----413	
Db 126856	AGTGACACCTGAGCTGGGACAGAAAGGCGTTTCTATTCTTCTTGAAAGCAGTGGCC	126915
OY 413	-----413	
Db 126916	TGTATCTTAAGATTTAAGAGACTAGGACAGATCAAGTTTACAGGCTGCTGACCTAT	126975
OY 413	-----413	
Db 126976	ATAGAAATTAATGTTTTCTCATGGAGGAAAGTTTTTTCCACAAAGCCTGAGGGTGCT	127035
OY 413	-----413	
Db 127036	CTTTCATCCATGGCACTTACAGACATGAGAGGACAGAACTGACACCCATTAATTCGCTTG	127095
OY 413	-----413	
Db 127096	ACCTCCACAGTGCAGGTGACTTGCACTACACAGAAACACATAAGTAATGTATT	127155
OY 413	-----413	
Db 127156	TAAAGACTCGAAAGGCTCAGGAAACAGTCTCAGCATGGGGGGGGGGGGGAG	127215
OY 413	-----413	
Db 127216	CATCCTTAGAGGTGGAGACAGACTAGCATACAGTCTTCCAGAAAGACGAAAGGGT	127275
OY 414	-----GlnThrProLeuGlyP419	
Db 127276	GGGTGACCTCTCCACACTGTGGCTGACGTCGAGGTCTCTTTGGACAGTCTCTTTGGCC	127335
OY 419	roAspTyrLysGlyCysLeuIleArgGlnAspProSerGlyAlaPheSerLeuValGlyL439	
Db 127336	CCGACTACAAAGGGCGGCTCATTCGGCCAGGACCCGACGGGGCTTCTCCGTGGTTGGCC	127395
OY 439	eulSerGlnProHisArgSerLeuArgGluLeuLeuAlaIaCysTrpAsnSerGlyLeuA459	
Db 127396	TCAGCAGACCCCAAGAAAGCTCGGGGAGCGTCTTCCAGCCCTGCGGAATCTCGGGTGC	127455
OY 459	IrgValAspGlyAlaAlaLeuTyrIleuThrSerCysAlaProArgProIys-----476	
Db 127456	GAGTAGAGGCGTGGCCCTTAACCTTAACATCTCTGCTGCCCTCCAGACCCAAAGGGTGAGT	127515
OY 476	-----476	
Db 127516	CAGTTCTCCCTAACCAAGCCGCGAGACTTTGGGTGGTAGAGAAACCAAGACCATGG	127575
OY 476	-----476	
Db 127576	CGATTAGACAGAGAACACGACTTTTGGCACAGTCTTCAGACATCAAGTCAGTATTATTTT	127635
OY 476	-----476	
Db 127636	CAGTTGGCGCGGCTTCTTCCATGGGTGAATGCCACTTAAGCGAACCCTGCTGGGACA	127695
OY 476	-----476	
Db 127696	CATCTCCCTCTGTGGTGTCTGGGCTTAACAAGACACAGGCGGCGCTGGAGGAAACCCAGT	127755

[illegible]

QY	591	-----	591
Db	128835	CCCTGGCCTCTCTCCACACCCTGCTACTACGTGATGCCTCTGATACCTTAAGATCATCATC	128894
QY	591	-----	591
Db	128895	CCCGTAGCACCCACCTACATCCAGAGGCCACACCGGGGACACGATCATCTTCCACC	128954
QY	591	-----	591
Db	128955	TTGACGACTAATTGCTATGCAACATTCCTCCACCCACCCCGGTGCCAACCCTCTCTGA	129014
QY	591	-----	591
Db	129015	ATCTTGCGTGTATCTATTCTAGTTAGCTGATAGAAACAGAGTCCAGCCTGTGCCAAGA	129074
QY	591	-----	591
Db	129075	GCAGAGGAGAAATGGGGGAGGCGCTGCTGAGTACCTGTCACCTGANTACACACACCCC	129134
QY	592	-----	604
Db	129135	ACTCTCCCCCTGTCTCCCGCAGGCATATGTCGAGGAATTGTGATCTAGAGCAATTCAC	129194
QY	605	MetTyrLeuArgLysArgGlyHisLeuValSerAlaSerThrPylsLeuGlnValThrLys	624
Db	129195	ATGTACTGTCGCGAAGCGTGGCCACCTGCTGTCACCCAGCTGGAACCTGCAAGGTGACCAAG	129254
QY	625	GlnLeuAlaTyrAlaLeuAsnTyr-----	632
Db	129255	CAGCTGGCATATGCCCCCTTAACCTACTTGTTGATGTCCCTGTGCCATTAAGAGAGTTCCT	129314
QY	632	-----	632
Db	129315	CTATGAGTGGGAAATGAGAGTAGAGTTCTGTGATTTGGATAGAGCTTGTTAAGGACACA	129374
QY	632	-----	632
Db	129375	AGCAACGTTCCAGATGACGTGAGCAGGACAGGCTCAGAGGGCTGTAAGTCTCAAGGGAGACA	129434
QY	632	-----	632
Db	129435	CTGTGAACCATGGGTGCTGTAGGAGACAGTAGTAGAGAGAGGCCAGGCTGAGGGGTGG	129494
QY	632	-----	632
Db	129495	CGGGGTGTAAGATGGAACCTGCTAGAGAAGAGGGGTGCTTGAATGGCGGATGTGGGG	129554
QY	632	-----	632
Db	129555	CTTTAACTGTCCTTAAGTCTGAGGAAGGTAGAGAAAGGGGTGTGTGCTTAAGATGGG	129614
QY	632	-----	632
Db	129615	TATGTCAGTGTAAAGGGAGGTTAAGGTTTACAAGACATGACTTACAGGGCTACAGTGGT	129674
QY	632	-----	632
Db	129675	TGCCAAGCCTGTTGTGTGTAGGACGACATCAGATAGGCTTTGAAGGATGACGAGAGTTT	129734
QY	633	-----	640
Db	129735	GCTCACAACTACTTCCCTGTTCCCTCTGTAACACTTAGGAGGACAAAGGCTTCTCTCA	129794
QY	640	sgLysAsnValSerAlaArgLysValLeuLeuAlaArgGlnGlyGlyAspGlyAsnProP	660
Db	129795	CGGGAAGCTTCCACACGGGAAGGTGCTCTGCTGCTGTAGGGGGGTGATGGGAATCCACC	129854
QY	660	oPheIleLysLeuSerAspProGlyValSerProThrValLeuSerLeuGlnMet-----	678
Db	129855	TTTATTAACTGATGATCTGTGTGTGTCAGTCCCACTGCTGGAGCCGTAAG- TGAGT	129913
QY	678	-----	678

Db 129914 ACAGGCTACCTCACCCTATACCCCGTGTCTGACGACACAGCTGTGCACAGAGGCT 129973
 Oy 678 -----
 Db 129974 GCACCTGGCAAGGGACCTTTGTGCAAGGTGACCATGACAGTTCTCCTACCTCCACGTT 130033
 Oy 679 -----LeuThrAspArgGlyLeuProTyrValAlaProGluCysLeuGlnGluAlaGlnThr 696
 Db 130034 TCCAGTGGCTCACCGACAGATATACCTGGGTGGCCCGCCGAAATGTCTCCAGAGGCTCACAG 130093
 Oy 696 rLeuCysLeuGluAlaAspLysTyrGlyPheGlyAlaThrThrTyrGluValPheGlnAr 716
 Db 130094 ACTGGCTGTGGAGCTGACAGAGTGGGCTTTGAGAGCCACACGATGGGAGGTTCACCGG 130153
 Oy 716 gGlyProAlaHisIleThrSerLeuGluProAla-Lys----- 728
 Db 130154 GGGACCGCCACATCACCTGGCTGGAGCCCGCCAAAGTCAGAACCGGTGTCTCGGCT 130213
 Oy 728 -----
 Db 130214 TCTGGCCCGGTGAGGGAGGGGATATCGTAGCGTAAATCGGAGCTGAGCA 130273
 Oy 729 -----LysLeuLysPheTyrGluAspGlnGly 738
 Db 130274 TCGCAACACCGAAGTGAACCTCGTCCCTTAGAAGCTGAAGTTCTATAGAGGACAGGAGAC 130333
 Oy 738 LLeuLysProAlaLeuLysTyrThrGluLeuAlaGlyLeuIleThrGlnCysMetAlaTyrA 758
 Db 130334 AGCTGCGCGCTCTCAATGAGACAGAACTGGCGGAGCTTATCACACAGCTGATGGCTTAG 130393
 Oy 758 sProGlyLysArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIleThrS 778
 Db 130394 ATCTGGCCCGCGCCCTCTCTCCGAGCTATCTCAGAGACCTCAACGGCTCATTAACAT 130453
 Oy 778 er----- 778
 Db 130454 CAGGTGCGGAGAGGGGGGGCGGCAAGGGCGGAGAGGCTCCCTGAGCTGGGTCCAGAA 130513
 Oy 778 ----- 778
 Db 130514 GTTGTGGTTCAGCAGTCACGCGTGAAGACTAGACAGCTGTGCACAAACCTAAGACCC 130573
 Oy 778 ----- 778
 Db 130574 GATCTCATTGTTTTCAGGTGAGCCAGGTTGGTGTTCATGATGAGCCCAACTAG 130633
 Oy 778 ----- 778
 Db 130634 CATCTAGGCTGATTTCTCCCTGCTTACATCATGAGTGGGTGGATTGGTTTATT 130693
 Oy 778 ----- 778
 Db 130694 TGTATTATTTAATTTTGTGCAAGATGTTCTCATCTGAGGATTCAGATCAGAGACT 130753
 Oy 778 ----- 778
 Db 130754 GGCCTACATGATGCTGTCTCTCGGGTCCCTTTGTCTGTAGCCGTTGGGTCCCA 130813
 Oy 778 ----- 778
 Db 130814 CAGAGTGTGCTTTCCCGGACAGTGTGCTTCATCCCATTAAGTACTGTGGCG 130873
 Oy 778 ----- 778
 Db 130874 TCTCATACGCACAGAGAGAGAGGTTCAGGTGAGTTTACACACCAATTCACTTTGGG 130933
 Oy 779 -----AspTyrG 781
 Db 130934 ATTGAGCAGAGACTTACCACTTACCCACTCTCTCAGGTGATCCCGCCCTAGATTAGC 130993
 Oy 781 LLeuLeuSerAspProThrProGlyIleProSerProArgAspGluLeuCysValAlaG 801
 |||||||
 |||||||

Db 130994 AGTCTCTCAGAGCCCAACACCTGGCATCCCGAGTCTCTCGAGTGAAGCTGTGC---GCTG 131050
 Oy 801 LysAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGluGlnArgHisLeuLysTyrI 821
 Db 131051 GCGCGCAGCTCTATGCTGTGCGAGACCCCGCCATATTTCAGGAGACACACTTAAGTACA 131110
 Oy 821 LeuSerLeuGlyLys----- 826
 Db 131111 TCTCTTGGCGGCAAGGTGAGTGGCGGCGCATGTGGGGAGAACCTGGGTGGTGA 131170
 Oy 827 -----GlyAsnPheGlySerValGluLeu 834
 Db 131171 GGGTCAGGTGACACTGCCCTCTCATCTCCCAAGGGCACTTTGGCAGCTGGAGCTG 131230
 Oy 835 CysArgTyrAspProLeuLysAspAsnThrGlyProLeuValAlaValysGlnLeuGln 854
 Db 131231 TGGCGCTATGACCCCTGGGGGACAAATACGGGACCCCTGTGGCAGTGAACAGCTTACG 131290
 Oy 855 HisSerValProAspGlnGlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeu 874
 Db 131291 CACAGGGGCGCACAGACAGAGGAGCTTCCAGGGGAGATTTCAGATCTTAAGGCTCTG 131350
 Oy 875 HisSerAspPheIleValLysTyrArgGlyValSerTyrGlyPro-----4----- 889
 Db 131351 CACAGGCACTTATGCTCAAGTACCGGGAGTCACTATGGGCCAGTGAAGCGGACGAG 131410
 Oy 889 ----- 889
 Db 131411 CATCTCGGGAAAGGGTTGAGTCCCGCTTCAACCTTTCCCAAGTAAAGGCCCTGTGAACA 131470
 Oy 889 ----- 889
 Db 131471 AGTGCTTAACCTCCATGAATCCAGCTTCTATAGCTGAGGTGAACGAATCAATACAC 131530
 Oy 890 -----GlyArgGlnSerLeuArgLeu 896
 Db 131531 CCAGTAGAAGGTCTCGAGACTACTGTGCTGTCAGAGTGCAGACCTGGGGTTG 131590
 Oy 897 ValMetGluTyrLeuProSerGlyCysLeuArgAspLeuGlnArgHisArgGly--- 915
 Db 131591 GTGATGAGTACCTGCGCCAGGCGCTGCTCGAGACTCTTCGAGCGCCATTCGCGGGCGC 131650
 Oy 916 LeuHisThrAspArgLeuLeuLeuPheAlaTyrGlnIleCysLys----- 930
 Db 131651 CTGCACACGACCGCCCTACAGCTGTGTGCTTGGCAGATCTGCAAGGTGCGAGCATCGCG 131710
 Oy 930 ----- 930
 Db 131711 AGATCGGAGTGGGCTCTCTGCGGTCCGAGTCCCGGGCCACGTCGATCATGAGCTCTCT 131770
 Oy 931 -----GlyMetGluTyrLeuGlnLysAlaArgArgCysValHisArgAspLeuAlaAlaArgAs 949
 Db 131771 GCAGGGCATGATGATCTGGGTGCGCGCGCTCGTACACCGTGAACCTGGCTGCCGCA 131830
 Oy 949 nIleLeuValGluSerGlnAlaHisValLysIleAlaAspPheGlyLeuAlaLysLeu 969
 Db 131831 CATCTTGTGTGAGAGAGGAGCTCATGTGAAGATCGGGACTTGGGCTGCTAAGCTGTCT 131890
 Oy 969 uProLeuGlyLysAspTyrTyrValValArgLysProGlyGlnSerProIlePheThrP-- 988
 Db 131891 GCCCTTGGGAAAGACTACTAGCTGTGCGGAGCTTGGCCAAAGCCCATCTTTTGGTA 131950
 Oy 988 ----- 988
 Db 131951 CAAGACCAAGACTCCAGAGTCAAGCGCCACTAGCCCAACTCCGCACTTACAGCCTT 132010
 Oy 988 ----- 988
 Db 132011 AAACGGTCCAGACCCAGCTCCACCTCAAGTACTTTGCCCTGCAAAAGGCTTGGCCACA 132070
 Oy 988 ----- 988
 Db 132071 GCAGGTGAGCCCTATAGCCAGATCCACGAAAGCATGCTCAAACTCTAGAGATGTACC 132130

[illegible]

QY	1064	-----	1064
Db 133210	CGATTGACAGCAGCGGTCGCTGACTGAGGAACTACTTACAGAGATGAGGACGCTCAGCTTC		133269
QY	1064	-----	1064
Db 133270	AAGGCTTCAGAAATGTGCTTTTCACCTTTTGTGAGGCAAGGCTGCTATGAGCCTTG		133329
QY	1064	-----	1064
Db 133330	TCTGGCTGGAACCTGCACATCTCTCTGCTCAGCTTCTCTGAGACTGACAGCAAAATGT		133389
QY	1064	-----	1064
Db 133390	ATGTGCACACACACACCACCCCACTCAGTATTTCAATTTTCTTTATATTACTTA		133449
QY	1064	-----	1064
Db 133450	TTTATTTTTCAGTGTGAGATTGGAAGCCAGGGTCTTACTATGCTAGGCAAGCAGCG		133509
QY	1064	-----	1064
Db 133510	ACCCGTGTGCATGATTAATTCGGGGGATGGGGGTTGGGGAAGGAGGGCAGCGCTTCCA		133569
QY	1064	-----	1064
Db 133570	GGAGTGAATGGGGCCTATTGGAGGAGCCAGAGAGAGTTCATTCACAGTAGAGACATGGCT		133629
QY	1064	-----	1064
Db 133630	GTCGGCTGGGGTGAGTAGAGAGACAGATGTGCGGCAGATGGGGGATGTGTGCTGTC		133689
QY	1064	-----	1064
Db 133690	ACCCGTGCATTCGCGAGCCAGGCGAGAGTGTGCCACAGGCAAACTTACTTCTGCTTC		133749
QY	1064	-----	1064
Db 133750	AGAGATCCATTAAGAGCCATGAGAGACTGTGTGAGAGAGTAGGGGCCATTAACTGAGT		133809
QY	1065	-----	1067
Db 133810	GGAAACACGCCCTCTCAGCTCAGAGCACCCCCACCTTCTCCACACCAAGTTCAGGA	ValGlnG1	133869
QY	1067	uLeuMetGlnLeuCysTrpAlaProGluProHisAspArgProAlaIleAlaThrLeuSeu	1087
Db 133870	GCATCATCAGCTGTGCTGGGCGCCGCGCAGCGGCGCAGCCGCTTGGCAGCCCTGAG		133929
QY	1087	rProGlnLeuAspProLeuTrpArgGlyArgProGly	1099
Db 133930	CCCCCAGCTGAGCGCGCTGTGGGCTGGAAGACCCCGGA		133966
RESULT 10			
MMU71201			
LOCUS	MMU71201	10326 bp	DNA linear 1
DEFINITION	Mus musculus protein tyrosine kinase JAK3 gene, complete cds.		
ACCESSION	U71201		
VERSION	U71201.1	GI:1778748	
KEYWORDS			
SOURCE			
ORGANISM	Mus musculus		
	Mus musculus		
	Eukaryota; Metazoa; Chordata; Cranulata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
REFERENCE	1 (bases 1 to 10326)		
AUTHORS	Kumar, A., Toscani, A., Rane, S. and Reddy, E.P.		
TITLE	Structural organization and chromosomal mapping of JAK3 locus		
JOURNAL	Oncogene 13 (9), 2009-2014 (1996)		
MEDLINE	9708640		
PUBMED	8934548		
REFERENCE	2 (bases 1 to 10326)		
AUTHORS	Kumar, A. and Reddy, E.P.		
TITLE	Direct Submission		
	Submitted (18-SEP-1996)		
	Fels Institute, Temple University, 3307 N.		

FEATURES Broad St, Rm 339 AHB, Philadelphia, PA 19140, USA
Location/Qualifiers
Source 1..10326

/organism="Mus musculus"
/strain="129/J"
/db_xref="taxon:10090"
/chromosome="19"
/map="19p13.1-p13.2"
Join(92..287,441..564,740..851,963..1108,1317..1599,
2170..2289,2372..2529,2617..2728,3230..3413,3781..3902,
3998..4129,4596..4680,5022..5149,5642..5774,5905..6056,
6177..6327,6856..6995,7074..7260,7427..7551,7629..7801,
8524..8641,8907..9017,9707..10212)
/product="JAK3"
Join(104..287,441..564,740..851,963..1108,1317..1599,
2170..2289,2372..2529,2617..2728,3230..3413,3781..3902,
3998..4129,4596..4680,5022..5149,5642..5774,5905..6056,
6177..6327,6856..6995,7074..7260,7427..7551,7629..7801,
8524..8641,8907..9017,9707..10212)
/function="protein tyrosine kinase"
/codon_start=1
/product="JAK3"
/protein_id="A040917.1"
/db_xref="gi:1778749"

CDS

/translation="MAPSEETPLIPQSCSSSSSAGALHWLLEPPRGPPQRLSFS
FGDYLAEDCYRAKAGGLIPVHSLFALATEDFSCPPSHIFCIEDYDQVLYRL
RIFYPMFGLECHRGKRLDLSALDLHLVLEHFAQRSDLVSGRLPVGLSMKEOG
EFLSLVLDLAOMARQAPQPGELIKTVSKLPSLRDVLGGNFPYRRIRRTYV
LALRVSPARPTAHGOVYLDLERHPATETPEVGLQGOEERGLIRVAGDNGIS
MSSGDELFOETCDPEIDVSIKOPTGSGREHRLVYTRNDGHLLEPEFGLEAL
SFVALDGFRLICDSRHYECKEVAPRLLEBEALCHPITLDRAIHKLAASLPG
TYLRSPODVSFLITACVPTGPBQKGLRQDPGSAFVLSLSPTEACGCLQ
CWNSGLRVGALNLTSCAPRREKSNLIVRQCTAPAPGCSPPSCALTOLEFHT
IPDYSCHENLHGSEFTKIFGRREVDGETHDEVLTKVDSHRRMCSEFLEAAS
LMSQVYPLVLHGVCMAGDSIMVOEEFYGAIMYLRKRLHGLVSAKMLQVTPOLA
VALNLTEDKGLPHGNVSAKVLAREGGGNPFIKLDQGVSPVLSLMLTTPRPA
VAPECLQEAQTLCEADKWFGATMEVSGPAITISLEPAKKLKEVEQCOPLAK
OLVACODPAIFEEHRLKYSLIGKNGPSVETCRDPIIDNTPVLAVALQDSPPDO
RDFOREIOLIKALHSDFTVYKRGVSGRQSLRVKMETLPSGCLRDQLQHRALHT
DRLIFAMQICKMEYLARRCVHRDLARNLIVSESHVNIADFGALKLPLTKDY
VVRPGQSTIFVAPESLSDNIFESKQSVSGVLYELFTYCDKSCSPSAEPRMNG
PERREPRCLRELEAEGRRRLPRPTCTEVOELQVLAASRTTQGPSAP"

BASE COUNT 2088 a 3007 c 2880 g 2300 t 51 others
Alignment Scores:
Pred. No.: 5,3e-232
Score: 3935.00 Length: 10326
Percent Similarity: 32.79% Matches: 1059
Best Local Similarity: 32.60% Conservative: 6
Query Match: 67.15% Mismatches: 31
Indels: 2163
Gaps: 25

US-09-397-967-16 (1-109) x MMU71201 (1-10326)

QY 1 MetAlaProSerGluIuThrProLeuIleProGlnArgSerCysSerLeuSerSer 20
Db 104 ATGGCACTCCCAAGTAGAGACACCTCTGATCCCTGAGCCCTCTTGAGCCTGCATCC 163
QY 21 SerGluAlaGlyAlaLeuHisValLeuLeuProProArgGlyProGlyProProGlnArg 40
Db 164 TCAGAGGAGGAGCCCTCATGTCTCTTCCCGGGGAGCCTGGGCCCTCCACACGGA 223
QY 41 LeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysValArgAlaAlaAlaAla 60
Db 224 TTGTCATTCTCTTTGGGGACTACTGGCTGAGATTATGTGTGCGAGAGCTCCCAAGGCC 283
QY 61 Cys----- 61
Db 284 TGTGGTAGTGGTCCCTAGATGACCATGCCATGTAAAGCAACATCTATGAGAAGGG 343
QY 61 ----- 61

Db 344 ATGCTGGGGTTTGGAGGGGTGTGGAATTTTGTAAAGGAGATCATGGGATGACCCCGTG 403
QY 62 ----- GlyIleLeuProValIleThrIleSer 69
Db 404 TAGTTGACTTTGGTGAACATGATGTGCCCATCCCTCGAGGATCTGCTGTATTATATATCG 463
QY 70 LeuPheAlaLeuAlaIleThrGlnAspPheSerCysTyrPheProProSerHisIlePheCys 89
Db 464 CTTTGGCTGTGCGCCACTGAGGACTTCTCTGCTGTGTTCCCAAGCCACATCTTCTGCG 523
QY 90 IleGluAspValAspThrGlnValLeuValIleTyrArgLeuArg----- 103
Db 524 ATAGAGAGAGTGGAGACACTCAAGCTTGGTCTACAGGCTAGAGTAAGTAAGAACCCCTTG 583
QY 103 ----- 103
Db 584 ATTATCTTCAGAGGATGATATTCTATTCATCATAGTAGAGGCTTTGAAACAGAGAGAG 643
QY 103 ----- 103
Db 644 TTGGGCTGTTGTGTAGAGCCACACTGCACCTTCACTGACCCCACTTGGCTGTGGAGAC 703
QY 104 ----- PheTyrPheProAspTyrPheGly 111
Db 704 CAGACTGAGTGGGCCACCCCATCTCTGAACCCACAGACTTTTATTTCCCTGACTGTTGG 763
QY 111 YLeuGluThrCysHisArgPheGlyLeuArgLysAspLeuThrSerAlaIleLeuAspLe 131
Db 764 GCTGAGACATGTGACCGCTTTGGGCTGCGGAAGATTGTGACAGTCCATCTTGACTT 823
QY 131 uHisValLeuIleuHisLeuPheAla----- 139
Db 824 ACATGTTTAAACATCTCTTGTCTCAGGTTGTATCTTGGTNGAGAGCTCAACCCA 883
QY 139 ----- 139
Db 884 GGGATGGGGAGCCCAAGGTAAACAATCCAGACCCAGACGGAACCTCATTAAACAC 943
QY 140 ----- GlnHisArgSerAspLeuValSerGlyArgLeuProValGlyLe 154
Db 944 CTTTCTGTGGCCGACCCAGACGCGCACTGATGACGTGAGTGGGCGGCTCCGGTGGCGCT 1003
QY 154 uSerMetLysGluGlnGlnIleuPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAl 174
Db 1004 TAGCATGAGAGAGAGAGGAGAGATTCTGACCTGGCGCTGCTGACTGTGACCATGATGCG 1063
QY 174 aArgGluGlnAlaGlnArgProGlyIleuLeuLeuLysThrVal----- 188
Db 1064 TCGTAGCAGGCGCCAGCGCCAGAGAGAGCTGAGAGAGGCTAGGTAGGGCTCGGANA 1123
QY 188 ----- 188
Db 1124 CCCCTTGTGTCTTCTTGGCTGCGGATCTAGGCGCCCTATGCTTTTCCCTCCGGANA 1183
QY 188 ----- 188
Db 1184 CACATACCTTCTCCCAACTTGTGACACCACTCTCCGTCAGCTGCCCCGGAAGGT 1243
QY 188 ----- 188
Db 1244 CCCCTATGCTGTCTGTCTGTCTTGGGCCCCCTACACCCCTAAGAGCTCGGCNTGNN 1303
QY 189 ----- SerTyrLysAlaCysLeuProProSerLeuArgAspValIleGlnGly 205
Db 1304 CCCCCCCCCACAGATTAAAGAGCTGTGCGCCGACGCTGCGGATGTATCCAGGGGCC 1363
QY 205 LAspPheValIleThrArgArgArgIleArgArgThrValValLeuAlaLeuLeuProCysG 225
Db 1364 AGAAGCTTGTGACAGGAGCGCATCCGACAGCGCTGTGCTGCGCTGC--CGGTGTG 1421
QY 225 YArgLeuProGlyArgProTyrAlaLeuMetAlaLysTyrIleLeuAspLeuIleuArgL 245

Db	4714	CGCTCTCCTCAACCCCTGCTCTACCTGCATGCTCCTGTGATACCTTAAGATATCATCACCCGTA	4773
QY	591	-----	591
Db	4774	CGACCCACACTCATATCCAGAGGCCACACCGGGCACACCGATCCATTTCACACTTGCCAGC	4833
QY	591	-----	591
Db	4834	TATTGCTATGCACATTCCCTCCACCCACACCCCGTCGCAACCCCTCCTGAATCTTGGC	4893
QY	591	-----	591
Db	4894	TGTATCTATTGAGTTGAGCTGATAGGAACAAGAAAGTCCAGCCTGTGCACAAAGACAGAGAA	4953
QY	591	-----	591
Db	4954	AAGTGGGGGACAGGCCCTGCTGTACTGCTACCTGATCAGTATCACCACCCACCCACTCTCCCGCT	5013
QY	592	-----	592
Db	5014	GTCCCCAGGACATCATATGGTGGCAGCAAGAAATTTGTGTATCTATGAGACAAATTCATGATACCTGGCG	5073
QY	608	gLyArGlyHtSLeuValSerAlaSerTPLYsLeuGlnValThrlrYsgLnuAlaTyr	628
Db	5074	CAACCGTGGCCACCTGTGTGCTGACCCAGCACTGGAACCTGAGGTGACCAAGCAGCTGGCAT	5133
QY	628	fAlSLeuAsnTyrLeu-----	633
Db	5134	TGCCCTTAACCTACTTGGTGGAGTGTCCCTGTGTCCATTAAGAGGCTTCCTATTGGATGG	5193
QY	633	-----	633
Db	5194	GGAATGAGAGTAGTGGTTCTGTATTGTTGATAGAGCTGGTAGACACAAAGCAAGCAAGTT	5253
QY	633	-----	633
Db	5254	CCAGAAATGATGTAGCAGCAGGCTCAGAGGGCTGTAAAGTGTCAAGGAGAGACACTGTGAACC	5313
QY	633	-----	633
Db	5314	ATGTTGTCTAGGGAAGCAGTGAAGTGAAGGACCAAGGCTGGAGGGGGGTGGCGGGGTGTA	5373
QY	633	-----	633
Db	5374	GAGATGGAACCTGCTGAGAAAGAGGGGTGTCTTGAAATGGCGGATGTGGCTTTAACTGTCC	5433
QY	633	-----	633
Db	5434	CTAAAGTCATGGGAAGGTAGAGAAAGGGGTGTGTCTTTAAGGATGGGTATGGTCACTGT	5493
QY	633	-----	633
Db	5494	AAGGGAGGTTAAGGTTACAAAGACATGACTTCACGGGCTTACAGTGGTTGCCAAAGCCTG	5553
QY	633	-----	633
Db	5554	TGTGTGTAGGCAGCATCTCAGTAAGCTTTGAAGATGCAGAGAGAGTGTCTTCACAACATA	5613
QY	634	-----	634
Db	5614	CTTCCCTGTCCCTCCTGTAAACCTTAGAGAGGACAAAGGCCCTTCTTCACGGCACAGCTTC	5673
QY	644	fAlaTArgysValLeuLeuAlaArgGluGlyGlyAspGlyAsnProPheIleIysLe	664
Db	5674	AGCAGGGAAGTGTCTCTGGCTCTGGAGGGGGGTGATGGGAATCCACCTTCATTAAAGCT	5733
QY	664	userAspProGlyValSerProThValLeuSerLeuGlu-----	677
Db	5734	GAGTATCTCGGTGTCAGTCCACAGCTGTCTACACCTGGAAGTAGTACAGGTAAGCTTC	5793
QY	677	-----	677

```

Db 5794 ACCCACTATACCCCGTGTCTGCAGACCACAGCTCCTCACACAGGGCCTGCACCTGGAAAGG 5853
Qy 678 -----MetLeuThra 681
      |||
Db 5854 GACCCCTTGGCAAGAGTGCACATGACAGTNCCTACCTCCACAGTCTCCAGTCTCATCCG 5913
Qy 681 spargileProtpPvalAlaProgluLysLeuIngluAlaInThrLeuCysLeuGlu 701
      |||
Db 5914 ACAGAAATACCTGGGTGGCCCGCAATGTCTCCAGAGAGCTCAGACACTCGCTTGAGAG 5973
Qy 701 laasplystrpGlyPheGlyAlaIathrThrprgluValPheGluInArgGlyProAlaHis 721
      |||
Db 5974 CTGCAAGTGGGGCTTTGGAGCCACACAGTGGAGGTTCAGCGGGGGACCCGCCACA 6033
Qy 721 lethrSerLeuGluProAlaLys-----
      |||
Db 6034 TCACCTCGCTGGAGCCCGCCAAAGTCAGAAACCTGTGTGGGTCTGGGCGCCAGC 6093
Qy 728 -----
      |||
Db 6094 GTGAAGGAGCGGGATATCTGTAGGGTGGTTAAATCTGAGCCGACATCGCAACAC 6153
Qy 729 -----LysLeuLysPheTyrgLuspgInglyLeuProA 741
      |||
Db 6154 CGAAGTGAACCTCGCTCCCTTAGAAGCTGAAGTTCTATGAGGACGAGGACAGCTGCCG 6213
Qy 741 laLeuLysTrpThrGluLeuAlaGlyLeuIleThrInCysMetAlaTyrsProGlyA 761
      |||
Db 6214 CTCTCAATATGACAGAACATGGCGGACCTTATCACACAGTGCATGGCTATATCTGGCC 6273
Qy 761 rgArpProSerPheArpGluAlaIleLeuArpAspLeuAsnGlyLeuIleThrSer 778
      |||
Db 6274 GGGGCCCTCTCTCCGAGTATCTCTCAGAGACCTCAAGGGCTATTCATCAGTGGCG 6333
Qy 778 -----
      |||
Db 6334 GAGAGGGGGGGGCGAGGGGCGAGAGGGTCCCTGAGCTGGGTCCAGAAAGTTGCTCG 6393
Qy 778 -----
      |||
Db 6394 TTCTAGACAGTCAAGCTGAACCTAAGACAGCTGCTGCAACAATCTAAGACCGATCTGCAT 6453
Qy 778 -----
      |||
Db 6454 TTGTTTGCAGTGTAGCCAGGTGGTGTATTACAATGTAGCCAAACTAGCATCTCAGG 6513
Qy 778 -----
      |||
Db 6514 CTGGATTTCNCCTGCTTACCATCATGAGGTGCTGGGATTGTTTATTATTCT 6573
Qy 778 -----
      |||
Db 6574 TAAATTTTGCACAGATGTCTCATCTAGATATCCAGTACACAGAGAGTCACTG 6633
Qy 778 -----
      |||
Db 6634 CATGGCTGTCTCTCGGTTCCCTTTGTCTAGCCGTGGGATCCACAGAGTGTGC 6693
Qy 778 -----
      |||
Db 6694 CTTTCCCGGACAGTGGTCTTCATTTCCCATAGTACTTGTGGCGTCTCATAGCGA 6753
Qy 778 -----
      |||
Db 6754 CACAGAGAGCAAGTTCAGGTGAGTTACACACAAATTCAGTTTGGGATTGAGCAGAG 6813
Qy 779 -----AspTyrgLusLeuLeuSerA 785
      |||
Db 6814 ACTTACACTTACCCCACTTCTCAGCTGATCCCGGCTAGAGTATAGAGACTCTCTCAG 6873
Qy 785 sspProthriProgluIleProSerProArpAspGluLeuCysValAlaGlyAlaGlnLeu 805
      |||
Db 6874 ACCCAACACCTGGGATCCCGAGTCTCGAGATGAGTGTGCTGCTGCTGCTGCTGCT 6930
      |||
Qy 805 yValAcysGlnAspProAlaIlePheGluGluArgHisLeuLysTyrlIleSerLeuLeu 825
      |||
Db 6931 ATGCCTGGCAGAGACCCCGCATATTGAGAGAGACACACTTAAGTACATCTTGTGCTGG 6990
Qy 825 Lys-----
      |||
Db 6991 GCAAGGTGATGGGGGGCATGTGGGGAGAACGTGGGTGGTGGATGGGTCAAGTGA 7050
Qy 827 -----GlyAsnPheGlySerValGluLeuCysArgTyrsP 839
      |||
Db 7051 CACTGCCCTCTCATCTCCACAGAGGCAACTTGGCAGAGGTGAGCTGTGCCCTATGACC 7110
Qy 839 roLeuLysPaspAnThrGlyProLeuValAlaValLysGlnLeuGlnHisSerValProA 859
      |||
Db 7111 CCTG---GACATACGGGAGCCCTGTGGTGCAGTAAACAGCTACAGCACAGCGGCCAG 7167
Qy 859 spGlnInArpAspPheGlnArpGluIleGlnIleLeuLysAlaLeuHisSerAspPhe 879
      |||
Db 7168 ACCAGCAGAGGAGACTCCAGCGGAGATTCAATCTTAAGCTCTGCACAGCGACTTCA 7227
Qy 879 leValLysTyrsArgLysAlaSerTyrgLysPro-----
      |||
Db 7228 TCGTCAAGTACCGGGAGAGTACGCTATGGCCAGGTGAGCGGCCAGCANCCTGGGAAGCG 7287
Qy 889 -----
      |||
Db 7288 GTTCGAGGNNCGCTTACNCTTNNCCAGTAGGCCGTTGAACAGTCTGTAACTCCAT 7347
Qy 889 -----
      |||
Db 7348 GAATCCAGCTTCTATAGCTGAGAGTTGAACGAACAATACCAGTAGAGTGTCTGAGA 7407
Qy 890 -----GlyArgGlnSerLeuArgLeuValMetLysIleuProSer 903
      |||
Db 7408 GTCACGTGCTGTTCGCGAGTCCGACAGGCTCGGTTGGTGTAGAGTACCGCCAGC 7467
Qy 904 GlyCysLeuArpAspLeuGlnArgHisArg---GlyLeuHisThrAspArgLeu 922
      |||
Db 7468 GGTGCTGGAGACTTCTCTGACAGGCTATCGCGGCGCTTGACACACGACCGCTACTG 7527
Qy 923 LeuPheAlaTrpGlnIle-----
      |||
Db 7528 CTGTTCGCTTGGCAGATCTGCAGAGTGCAGATCGCGAGATCGGGATGGGCTCTGGG 7587
Qy 929 -----CysLysGlyMetGluTyrlLeuGly 936
      |||
Db 7588 GTCGGAGTCCCGGCCCAACGCTGATCATGCTCTCTGAG-GGCATGAGTACTGAGGT 7646
Qy 937 AlaArgArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGlnSerGluAla 956
      |||
Db 7647 GCGGCGCGCTCGTACACCGTGAACCTGCTGCCGCAACATCTTGTGGAGAGCGAGCT 7706
Qy 957 HisValLysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAspTyrlTy 976
      |||
Db 7707 CATGTGAAGATCGGGAGCTTGTGGCTCTGCTTACAGCTGTGCCCCCGGAAAGGATACTAC 7766
Qy 977 ValValArgGluProgluIleGlnSerProIlePheTrpTyrla-----
      |||
Db 7767 GTGTCCGGAGAGCTGGCCAAAGCCCATTTTGTGTACAA-GACCAAGACTCCAGAGCT 7825
Qy 990 -----
      |||
Db 7826 CACGCCACTAGGCCCAACTCCCGACTTACAGCTTAACCGGTCCAGACCCAGCTC 7885
Qy 990 -----
      |||
Db 7886 CACTTCAGTACGCCCTGCTCGCAAGGCTTGCACAGCAGGGTCAAGCCTATAGCAC 7945
Qy 990 -----
      |||
Db 7946 GATCCAGAAAGCAATGCTCAACAATCTACAGATGTACCCCTCTCAACTTAGTCAGACTA 8005
      |||

```

QY 990 ----- 990
 Db 8006 CCCACCACCCACAGAGATGATGCTCAACNGGTTTACCAAGGCCCATTTGTTCTGGC 8065
 QY 990 ----- 990
 Db 8066 TCCATCTCTTGCATGACAGTCTGCMAAGCCGTCACAGAGTCTCTTTTCTAGTCTGCTGTC 8125
 QY 990 ----- 990
 Db 8126 CCGAGCCCTGCTTCAACGCTCTATGGGCTGCTGTCTTACGCTATTCTACTTCTCGCA 8185
 QY 990 ----- 990
 Db 8186 TCTTACTCTTCTCTCTGCTGCTGCTCTCCCTTATGACACCCATGCTAGTCTGCC 8245
 QY 990 ----- 990
 Db 8246 AGGCNCTTTTCCAGTCTGCTGCTCTCTCAGTCCGGTCTGACCCACTCAAGCCCGCTCTAA 8305
 QY 990 ----- 990
 Db 8306 CACTTAGAGTCCCTCTCAACCTGCGCCGACAGGTACACAGAGCTCTGAGTGC 8365
 QY 990 ----- 990
 Db 8366 CTTTCTGGGGCCGCTGCTAGTAGAAGTCTCTGTCACGCCCTTCTCTACCCAGAC 8425
 QY 990 ----- 990
 Db 8426 TGGCATTACTTACATTGAGAGACTCCCTCCCTGCGCCACTGTTTACACCCCTGCCACTC 8485
 QY 991 ----- 991
 Db 8486 TCTTACTGCGCCACACACACCTCTCTCCGCTCTCAGTATGCGCCGAGCTCCATCT 8545
 QY 996 Asphanillepseferarjgluseraspvaltrpserphelvalyalleutyrgluleu 1015
 Db 8546 GACAAACATCTTCCCGCCCAATCTGACGTGTGGAGCTTGGAGTGTGGTGTGAGAGCTC 8605
 QY 1016 PhehtrlyrcysasplysSerCysSerProSerAla----- 1027
 Db 8606 TTTCACTTACGTGCGACAGAGCTGACGCCCATCCGCTGTGCTGCTGCCGCCCATCCCTGCG 8665
 QY 1027 ----- 1027
 Db 8666 TCCCTCCCTTGTGATCTCCAAATCCCTTCTGACCTTACNCCCTATCTGACCCAGCC 8725
 QY 1027 ----- 1027
 Db 8726 TTTCTCTGACCTTCAGATCATCTCTGACCTGACCTCTCCCTCCGGAATCCCANCCCTT 8785
 QY 1027 ----- 1027
 Db 8786 CTTCTCATCTCCAGATCCCTTCTTGTGACCCAGCCCTTCCCTGANNCCANCCCTTCT 8845
 QY 1027 ----- 1027
 Db 8846 CATGACNTGANCNCCCTTCTGACCTTAACAGTCTNCCCTTACTTCTCTGCTCCACA 8905
 QY 1028 -GlupheleuAargmetGelyProgluaArgLuglyProProLeuCyArgLeuLeu 1047
 Db 8906 GGGATTCCTGCGCATATGGGGCTGAGCGGTGAAGAGACCCCGCTGCGCGCTCTCTGA 8965
 QY 1047 uleuLeuAlaLuglyArgArgLeuProProProthryCysProthryGluVal----- 1065
 Db 8966 GCGGCGTGGAGAGGGCGCGCTCCACACCTCCACCTGCGCCCGAGGT- GAGGA 9024
 QY 1065 ----- 1065
 Db 9025 AGGAGACTCAAGTTTCTCAGTTCAATTTCTGATTGACAGCAGCGGCGCTGACTGGGA 9084
 QY 1065 ----- 1065

Db 9085 ATACTTACAGATGAGGAGCGCTCAGCTTCAAGGGCTTCAAGAAATGTGTTTCACTTT 9144
 QY 1065 ----- 1065
 Db 9145 TGTGAGGCAAGGCTTGTCTATGTAGCTTGTGTGCTGTGAACTGACATCTCTGCGC 9204
 QY 1065 ----- 1065
 Db 9205 TCAATTTCTCTGAGCTGAGACAAATGTATGTGCCACACACACCCCACTCC 9264
 QY 1065 ----- 1065
 Db 9265 ATATTCAATTTTCTTTTATTTATTTACTTATTTATTTTAAAGTCTGAGATTTGAAG 9324
 QY 1065 ----- 1065
 Db 9325 CCAAGGCTTACTCATGTATGACCAAGCAGCACCTGTGTATATTTATTCGGGGATAG 9384
 QY 1065 ----- 1065
 Db 9385 GGGTGGGGAAGGAGGCGCGGCTTCAAGAGTGAATGGCCCTATTGAGAGCCAGG 9444
 QY 1065 ----- 1065
 Db 9445 AGAGTCAATCCAGTAGAGACATGCTGTGCGCTGGGCTGAGTAGAGAGCAGATG 9504
 QY 1065 ----- 1065
 Db 9505 TCGNGGAGATGGGGATGTGTGCTGTGACCTTACCATCTGACGGCAGGAGAGT 9564
 QY 1065 ----- 1065
 Db 9565 GTTGCACAGGCACTAGTATTCTGTCTGACAGATCATAGAGAGCATGAGAGCTGTGG 9624
 QY 1065 ----- 1065
 Db 9625 AGAAGTAGANNNGCATTTAACTGAATGAACACACACTCTCAGCTCAGCATCCGCCA 9684
 QY 1066 ----- 1066
 Db 9685 CCTTGATCTTCCACACACAGGTTACAGAGCTCATGAGCTGTGCGTGGCGCCAG-CC 9743
 QY 1077 chiSparpProAlaPheAlaThrLeuSerProgluLeuAspProLeuTrpArglyar 1097
 Db 9744 GCGAGCGCGGACAGCTTGTGCGACCTGAGCCCGACCTGAGCGCTGTGGGTGGAAG 9803
 QY 1097 gpro 1098
 Db 9804 ACCC 9807
 RESULT 11
 AF034576
 LOCUS 3644 bp mRNA linear VRT 03-SEP-1998
 DEFINITION Gallus janus tyrosine kinase (JAK) mRNA, complete cds.
 ACCESSION AF034576
 VERSION AF034576.1 GI:2645986
 KEYWORDS
 SOURCE
 ORGANISM
 Gallus gallus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 Phasianidae; Gallus.
 REFERENCE
 1 (bases 1 to 3644)
 AUTHORS Sofer,L., Kampa,D. and Burnside,J.
 TITLE Molecular cloning of a chicken JAK homolog from activated T cells
 JOURNAL Gene 215 (1), 29-36 (1998)
 MEDLINE 9832716
 PUBMED 9666067
 REFERENCE
 2 (bases 1 to 3644)
 AUTHORS Sofer,L., Kampa,D. and Burnside,J.
 TITLE Direct Submission
 JOURNAL Submitted (14-NOV-1997) Animal Science, University of Delaware, 40

Oy	474	ArgProLySGluYSerAsnLeuIleValAlaArgArgLys-----Asn	489
Db	1516	CTGGCCCGGAGAAAGTCAACCTGGTATTCGGCAGTGGATCCACAGACCCCAAT	1572
Oy	490	ProAlaProAlaProGlyCysSerProSerGysAlaLeuThrGlnLeuSerPheHis	509
Db	1576	TCGCCCCCGCCCCCGCCAGCCCAAC-----CAGATGTCATTCCAC	1620
Oy	510	ThrIleProThrAspSerLeuGluTrpHisGlnAsnLeuGlnAspSerPheThrLys	529
Db	1621	AAAGTTCAGCCCCGAGAGCCATACGGGGGAGAGCCGTGGGGAGGGGCTCTTCACCCAC	1688
Oy	530	IlePheArgGlySerArgArgGluValAlaAspGlyLysThrHisAspSerGluValLeu	549
Db	1681	ATCTCAAAAGGATGATACGGGGAGGACAGAGAGATGACAGTTTTCACAGACCCCGTGGG	1740
Oy	550	LeuLysValMetAspSerArgHisArgAsnCysMetLysSerPheLeuGlnAlaAsp	569
Db	1741	CTCAAGAGTGATGACAGACGCCACCCCAAGTCCGAGTCCGATCTCTGGAGGCCGCCAC	1800
Oy	570	LeuMetSerGlnValSerTyrProHisLeuValLeuLeuHisGlyValCysMetAlaCly	589
Db	1801	ATCATGAGCGACGCTCCCAACACACCTGTCTGTGACGGCTCTACGCTCGGCAAG	1860
Oy	590	AspSerIleMetValGlnGlyPheValTyrLeuGlnValAlaIleAspMetTyrLeuArgLys	609
Db	1861	GACACATCATGTGTGACAGGATACATAGGACAGGCCCCCTGTGACTCTTACCTAAAGAG	1920
Oy	610	ArgGlyHis-----LeuValSerAlaSerTyrPlyLeuGlnValThrLysGlnLeuAla	627
Db	1921	AACCAACCGGAGGACAGGTGACACACCTGGAACTCGAGTGGCCAGACAGCTGGCA	1980
Oy	628	TyrAlaLeuAsnTyrLeuGlnAspLysGlyLeuProHisGlyAsnValSerLysArgLys	647
Db	1981	TATGACATCACTACTTGTGGAGGATAGAAATATACCAGGCAAGCTGTCTGTAAAGAG	2040
Oy	648	ValLeuLeuAlaArgGlyLys-----GlyAspGlyAsnProPheIleLysLeuSer	665
Db	2041	GTCGCTGACCCCGGAGGGGAGCGCGCCACAGCACGCCCCCTCATCAAGCTCAAC	2100
Oy	666	AspProGlyValSerProThrValLeuSerLeuGlnMetLeuThrAspArgIleProTrp	685
Db	2101	GACCCCGGGGTCAACATCACCCCTCTGTGGCAGAGATGGCTGTGAGAGGCAATCCCTTGG	2160
Oy	686	ValAlaProGlyCysLeuGlnGlnAlaGlnThrLeuCysLeuGlnAlaAspLysTrpGly	705
Db	2161	GTCGCCCCCGAGTCCGTGACGACGCCCCAGAGGCTGGCCCTGCAGCTGCACAAATGGGGC	2220
Oy	706	PheGlyLysThrTrpTrpGlyValAlaPheGlnAlaGlyLysProAlaHisIleThrSerLeuGln	725
Db	2221	TTTGGGGCAACCTTATGGGAGATCTTCACGTGGGGCAACATGCTCTGACGCTACTGGAG	2280
Oy	726	ProAlaLysLysLeuLysPheTyrGlnAspArgGlnGlnLeuProAlaLeuLysTrpThr	745
Db	2281	CCACAGAGAAGGTGCGAGTCTTACGACACCCCTGTGCAGCTGCCGCGCCGCTGGAGC	2340
Oy	746	GluLeuAlaGlyLeuIleThrGlnCysMetAlaTyrAspProGlyAlaArgProSerPhe	765
Db	2341	GAGCTGGCCCGCTCATTCCTCTAGTGCATAC--TACGGGCCAGAGAGCGGCGCTGCTTC	2397
Oy	766	ArgAlaIleLeuArgAspLeuAsnGlyLeuIleThrSerAspTyrGlnLeuLeuSerAsp	785
Db	2398	CGTGCATCATCCCGCAGCATCAACAGCCATCTCTCTCGACATCGACAGCTGCTTCAGAG	2457
Oy	786	ProThrProGlyIleProSerProArgAspGluLeuCysValAlaGlyAlaGlnLeuTyr	805
Db	2458	CTGTGACCCCGCATGTGACGCTGGCG--GAGAGCTGTGCGGGGATAGACACAGTGGCG	2514
Oy	806	AlaArgGlnAspProAlaIlePheGlnGlnAlaGlnHisLeuLysTyrIleSerLeuLeuGly	825
Db	2515	CGGGGGCAGCGCCGCTCAGTTCCAGGGAGGAGCACCTCAAGTCAATCTACATCTGTCGGA	2574

OY	826	LysGIAlsnPnPhgIysSerValGlnLeucysArgTrpAspProLeuGlyAspAsnThrGly	845
Db	2575	AAGGGCAACTTTGGACCGCTGGACCTGTGGCCGTACAGACCGCGTGGTGAGACAGCAGCGT	2634
OY	846	ProLeuValAlaValLysGlnLeuGlnHisSerValProAspGlnGlnArgAspPheGln	865
Db	2635	GAGCTGTGGCCCGCGAAGAAGATTCACAGAGATTTCGGGCCAAAGAGACGTGCAGACTTTGAG	2694
OY	866	ArgGluIleGlnIleLeuLysAlaLeuHisSerAspPheIleValLysTrpArgGlyVal	885
Db	2695	AGGAGATCCAGATCTCGACCTGCTCACACACACTTCATCTGCACAATGACCGGGCGTCC	2754
OY	886	SerTrpIleProGlyArgGlnSerLeuArgLeuValMetGlnLysTrpLeuProSerGlyCys	905
Db	2755	TGTACACAGCGGTGGGGCGGGGGCTGCGGTGTGTACAGTAATCTGCGCTGACGCGCTGC	2814
OY	906	LeuArgAspLeuLeuGlnArg-----HisArgGlyLeuHisThrAspArgLeuLeuLeu	923
Db	2815	CTTAGGAGACACTCTGCMAAAGAACACAGCACCGC---CTGAGACACCACGCTGCTCTC	2871
OY	924	PheAlaTrpGlnIleCysLysGlyMetGlnLysTrpLeuGlnLysAlaArgArgCysValHisArg	943
Db	2872	TACGCTGGCAGATCTTGCAAGGCAATGAGTAACCTGGGGCGGACGCTCGTGCACCGC	2931
OY	944	AspLeuAlaAlaArgAsnIleLeuValGlnSerGluAlaHisValLysIleAlaAspPhe	963
Db	2932	GACTTGGCCAGCAGGAACATCTGGTGGAGAGAGACGACCCACCATCAAGATCGGTGACTTC	2991
OY	964	GlyLeuAlaLysLeuLeuProLeuGlnLysAspTrpTyValValAlaArgGluProGlyGln	983
Db	2992	GGGCTGGCCAAAGCTGCTGCACAGCAAGAGACTACTGCTGTCGACGAGAGCCGGGAG	3051
OY	984	SerProIlePheTrpTyraIleProGlnSerLeuSerAspAsnIlePheSerArgGlnSer	1003
Db	3052	AGCCCCGTGTTCGTGATGCCACCCGAGTGGCTGGTGCACAGCTTCTTCCGACATCC	3111
OY	1004	AspValThrSerPheGlyValValLeuTyrgLysLeuPheThrTyrcysAspLysSerCys	1023
Db	3112	GACGTGTGAGCTTGGGGTGGTGTCTCTATGAGCTCTTCACCTACACATAAGACAGAG	3171
OY	1024	SerProSerAlaGlnPheLeuArgMetMetGlyProGlnArgGlnLysProProLeuCys	1043
Db	3172	ACCCCCTGGAGAGATTTCTCCATGATGAGGCCCTGGAADAAACCAACAGATCATCTGC	3231
OY	1044	ArgLeuLeuGlnLeuLeuAlaGlnGlyArgArgLeuProProProThrCysProThr	1063
Db	3232	CATTTCGTGGAGCTGCTGAAGACATCCCGCGGCTCCGGTCCCGCGGCTGCCCCATG	3291
OY	1064	GluValGlnGlnLeuMetGlnLeuCystrpAlaIleProGluProHisAspArgProAlaPhe	1083
Db	3292	GAGGTGTATCGATGATGCTGAGCTGGCTGGCGTTCCGCCCAAGTCCAGACCCACCTTC	3351
OY	1084	AlaThrSerProGlnLeuAspProLeuThrParGlyArg	1097
Db	3352	ACCGAGTTGGCAGCCAGGTTGAGGCACCTGCGGAGTGGCCGT	3393
RESULT 12			
AX203084			
LOCUS	AX203084	2914 bp	DNA linear PAT 30-AUG-2001
DEFINITION	Sequence 5 from Patent WO0152892.		
ACCESSION	AX203084		
VERSION	AX203084.1	GI:15392441	
KEYWORDS	human.		
SOURCE	Homo sapiens		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	Jak/stat pathway inhibitors and the uses thereof		
TITLE	Patent: WO 0152892-A 5 26-Jul-2001;		
JOURNAL	GENZYME CORPORATION (US)		
FEATURES	location/Qualifiers		

```

source          1. .2914
                  /organism="Homo sapiens"
                  /db_xref="taxon:9606"
BASE COUNT      566 a      907 c      809 g      632 t
ORIGIN
Alignment Scores:
Pred. No.:      2.85e-189      Length:      2914
Score:          3228.50      Matches:      614
Percent Similarity: 87.63%      Conservative: 45
Best Local Similarity: 81.65%      Mismatches: 90
Query Match:    53.09%      Indels:      3
                        Gaps:      3
US-09-397-967-16 (1-1099) x AX203084 (1-2914)
OY 350 AsperAargHISYRPhCysLysGluValAlaProProAtrLeuLugLugLAla 369
    |||.....|||.....|||.....|||.....|||.....|||.....|||
Db 5 GACTCCAGCACTTCTGTGCAGAGAGGTGGCAGCCGCGAGCTGTGAGAGAGTGGCC 64
OY 370 AspaValcysHISglProIleThrLeuaspPheAlaIleHISLysLeuLysAlaAlaGly 389
    |||.....|||.....|||.....|||.....|||.....|||.....|||
Db 65 GAGCAGTGCACAGGCCCATCATCTGGACTTTGGCATCAACAGCTCAAGACTGGGGCC 124
OY 390 SerLeuProGlyThrTyrIleLeuAargArgSerProGlnAspTyrAspSerPheLeuLeu 409
    |||.....|||.....|||.....|||.....|||.....|||.....|||
Db 125 TCACGTCCTGGCTCTATGTTCTCCGCGCAGCCCGCAGACTTTGACAGCTTCTCTC 184
OY 410 ThrAlaCysValGlnThrProLeuLysProAspTyrLysGlyLysLeuIleAargGlnAsp 429
    |||.....|||.....|||.....|||.....|||.....|||.....|||
Db 185 ACTGCTGTGTGCACAAACCCCTTGGCTGTGATATTAAGGGCTGCTCATCGCGGCGAGC 244
OY 430 ProSerGlyAlaPheSerLeuValGlyLeuSerGlnProHISArgSerLeuAargGlnLeu 449
    |||.....|||.....|||.....|||.....|||.....|||.....|||
Db 245 CCCACAGCACTTCTCTGTGTTGGCTTCACGCCAGCCCGCAGCAAGCTTCAGAGCTC 304
OY 450 LeuAlaAlaCysTrpAsnSerGlyLeuAargValAlaAspGlyAlaIleLeuTyrLeuThrSer 469
    |||.....|||.....|||.....|||.....|||.....|||.....|||
Db 305 CTGGCAGACTGCTGGATGGGGGCTGCACGTAAGATGGGTGGCAGTACCCTCACTTCC 364
OY 470 CysCysAlaProAargProLysGlyLysSerAsnLeuIleValAlaIArgArgGlyCysAsn 489
    |||.....|||.....|||.....|||.....|||.....|||.....|||
Db 365 TCTGTATCCCGAGACCCCAAGAAAAGTCCAACTGATGCTGTCCAGAGAGTACAGC 424
OY 490 ProAlaProAlaProGlyLysSerPro---SerCysCysAlaLeuThrGlnLeuSerPhe 508
    |||.....|||.....|||.....|||.....|||.....|||.....|||
Db 425 CCACCCCAATCACTCTGGTTGACGCCCAATCCCAATACCAGCTGAGTACATGATTT 484
OY 509 HISThrIleProThrAspSerLeuGlnTrpHISGlnAsnLeuGlyHISGlySerPheThr 528
    |||.....|||.....|||.....|||.....|||.....|||.....|||
Db 485 CACAAGATCCCTGTGTCAGAGCCGTGAGTGGCATGAGAACCTGGGCCATGGTCTTCACC 544
OY 529 LysIlePheAargGlySerArgArgGlnValAlaAspGlyGlnThrHISAspSerGlnVal 548
    |||.....|||.....|||.....|||.....|||.....|||.....|||
Db 545 AATATTACCGGGGCTGTGCTCATGAGTGTGTGATGGGAGAGGCCGGAAGACAGAGGTG 604
OY 549 LeuLeuLysValMetAspSerArgHISArgAsnCysMetGlnSerPheLeuGlnAlaAla 568
    |||.....|||.....|||.....|||.....|||.....|||.....|||
Db 605 CTCTCTAGAGTCTATGATGCGACAGCACAAGAACTGCATGAGTCACTTCGGAAGCAGCG 664
OY 569 SerLeuMetSerGlnValSerTyrProHISLeuValLeuLeuHISLysIleValCysMetAla 588
    |||.....|||.....|||.....|||.....|||.....|||.....|||
Db 665 AGCTTATGATGAGCCAAAGTGTGTACCGGCACTCTGCTGTCCACGCGCTGTCCAGGCT 724
OY 589 GlyAspSerIleMetValGlnGlnPheValTyrLeuGlnValAlaIleAspMetTyrLeuAarg 608
    |||.....|||.....|||.....|||.....|||.....|||.....|||
Db 725 GGAAGACAGCACCATGGGAGGAGATTGTACACCTGGGGGCCATAGCATGTATCTGGCA 784
OY 609 LysAargGlyHISLeuValSerAlaSerTyrLysLeuGlnValThrLysGlnLeuAlaTyr 628
    |||.....|||.....|||.....|||.....|||.....|||.....|||
Db 785 AAAGGTGCGCAACCTGGTGCAGCAGCAGCTGCAAGCTGAGTGTCAAAAGCTGGGCTAC 844
OY 629 AlaLeuAsnTyrLeuGlnAspLysGlyLeuProHISLysValSerAlaAargLysVal 648

```

```

Db 845 GCCCTTAATATGTGAGAGACAAAGGCCCTGCCCATGGCAATGTCTTCGCGAAGGTG 904
OY 649 LeuLeuAlaAargGlyGlyLysAsnProProPheIleLysLeuSerAspProGly 668
    |||.....|||.....|||.....|||.....|||.....|||.....|||
Db 905 CTCTGAGCTCGGAGGAGGCTGTATGGGAGCCCGCTTATCAAGCTGTAGTACCCTGGG 964
OY 669 ValSerProThrAlleLeuSerLeuGlnMetLeuThrAspArgIleProTyrValAlaPro 688
    |||.....|||.....|||.....|||.....|||.....|||.....|||
Db 965 GTACAGCCCGCTGTGTAAGCTGAGATGCTTACGACAGATGCCCTGGGTGGGCCCC 1024
OY 689 GlnCysLeuGlnGlnAlaGlnThrLeuCysLeuGlnAlaAspLysTyrPheGlyAla 708
    |||.....|||.....|||.....|||.....|||.....|||.....|||
Db 1025 GAGTGTCTCGGGAGGGCGAGACACTTAGCTTGAAGCTGACAACTGGGCTTCGGCGCC 1084
OY 709 ThrThrProGlnValPheGlnAargGlyProAlaHISIleThrSerLeuGlnPheAlaTyr 728
    |||.....|||.....|||.....|||.....|||.....|||.....|||
Db 1085 ACAGTCTGGAGAGTGTATTAGTGGCTGCACATGCCCATCAAGTCCCTGATCTGTAAG 1144
OY 729 LysLeuLysPheTyrGlnAspGlnGlnLeuProAlaLeuLysTyrThrGlnLeuAla 748
    |||.....|||.....|||.....|||.....|||.....|||.....|||
Db 1145 AAATCCCAATTTTATGAGACCGGACAGCAGCTGCGGCGCCCAAGTGGACAGAGCTGGCC 1204
OY 749 GlyLeuIleThrGlnCysMetAlaTyrAspProGlyAargArgProSerPheAargAlaIle 768
    |||.....|||.....|||.....|||.....|||.....|||.....|||
Db 1205 CTGCTGATTCACAGCTCATGAGCTTATGAGCCGCTCCAGAGGCTCTCTCGAGCCGTC 1264
OY 769 LeuAargAspLeuAsnGlyLeuIleThrSerAspTyrGlnLeuLeuSerAspProThrPro 788
    |||.....|||.....|||.....|||.....|||.....|||.....|||
Db 1265 ATTGCTAATCTCAATACCTTCATCTCTTCAAGCATAGAGTCTCTTCAGACCCACACT 1324
OY 789 GlyIleProSerProAargAspGlyLeuCysValAlaGlyAlaGlnLeuTyrAlaCysGln 808
    |||.....|||.....|||.....|||.....|||.....|||.....|||
Db 1325 GTCGCTGAGCACTCTGTGATGGCTGTG---AATGTCGCCAGCTTATGCTGCGCAA 1381
OY 809 AsproAlaIlePheGlnGlnAargHISLeuLysTyrIleSerLeuLeuGlyLysGlyAsn 828
    |||.....|||.....|||.....|||.....|||.....|||.....|||
Db 1382 GACCCACGATCTTCAGAGAGACACCTCAAGTACATCTCACAGCTGGGCAAGGGCAAC 1441
OY 829 PheGlySerValGlnLeuCysAargTyrAspProLeuGlyAspAsnThrGlyProLeuVal 848
    |||.....|||.....|||.....|||.....|||.....|||.....|||
Db 1442 TTGGCAGCGTGGAGCTGTGCGCTATGATGACCCGCTACCCCAATACAGGTGCCCTGTG 1501
OY 849 AlaValLysGlnLeuGlnHISerValProAspGlnGlnAargAspPheGlnArgGlyIle 868
    |||.....|||.....|||.....|||.....|||.....|||.....|||
Db 1502 GCCGTGAACAGCTGCACACAGCGGCGCAGACCAACAGAGGACTTTCAGCGGAGATT 1561
OY 869 GlnIleLeuLysAlaLeuHISerAspPheIleValLysTyrAargGlyValSerTyrGly 888
    |||.....|||.....|||.....|||.....|||.....|||.....|||
Db 1562 CAGATCCCTCAAGAGCACTGCACAGTATTTCAATGTCAAGTATCGTGTGTCAGCTATGGC 1621
OY 889 ProGlyAargGlnSerLeuAargValMetGlnTyrLeuProSerGlyCysLeuAargAsp 908
    |||.....|||.....|||.....|||.....|||.....|||.....|||
Db 1622 CCGGGCCGCGCCAGACACTCGCTGCTCATGAGTACCTGCGGAGAGGCTGCTGGCGAC 1681
OY 909 LeuLeuGlnAargHISArgGly---LeuHISThrAspAtrLeuLeuPheAlaTyrGln 927
    |||.....|||.....|||.....|||.....|||.....|||.....|||
Db 1682 TTCTCTGACAGCGGACCGCGCGCTGATGCGAGCGGCTCTCTCTATCTCTCGAG 1741
OY 928 IleCysLysGlyMetGlnTyrLeuGlyAlaAargArgCysValHISArgAspLeuAlaAla 947
    |||.....|||.....|||.....|||.....|||.....|||.....|||
Db 1742 AACTGCAGAGGCGATGAGTACTGAGGCTCCCGCGCTCTGTCAGCCGAGCTGGCGGCC 1801
OY 948 ArgAsnIleLeuValGlnSerGlnAlaHISValLysIleAlaAspPheGlyLeuAlaLys 967
    |||.....|||.....|||.....|||.....|||.....|||.....|||
Db 1802 CGAATCATCTGTGTGAAGAGCGAGGACACATGCAAGATCGCTGACTTGGCTACCTAAG 1861
OY 968 LeuLeuProLeuGlyLysAspTyrTyrValAlaAargGlnProGlyGlnSerProIlePhe 987
    |||.....|||.....|||.....|||.....|||.....|||.....|||
Db 1862 CTGCTGCCCTTGACAAAGACTACAGTACGTGTCCGAGCGCAGGCGCAAGGCCCATTTTC 1921
OY 988 TrpTyrAlaProGlnLeuSerLeuSerAsnIlePheSerAargGlnSerAspValTyrPhe 1007
    |||.....|||.....|||.....|||.....|||.....|||.....|||

```


QY 140 LnhIaSerSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLysGlnG 160
 |||||
 Db 1460 AGCAGCGCAGTGTGAGTGTGGGGCCCTCCCTGGGGCTGAGTCTCAAGAGCAGG 1519
 QY 160 LysGluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGlnAlaGln 180
 |||||
 Db 1520 GTGAGTGTCTCAAGCTGCGCCCTGTTGGACCTGCGCCGATGGCGCAGACAGCCCAAGC 1579
 QY 180 rPProGlyGluLeuLeuLysThrVal----- 188
 |||||
 Db 1580 GCGCGGGGAGACCTCTGTAAGACTCTCAGGTAGAGACCACCAAGCGTGTGGGACGGGCTCT 1639
 QY 188 ----- 188
 Db 1640 GTTTGGAGTGTAGCAACGTGGGCTCATCGGGGGTTGGCGGGCTCCACCATGAGATTTC 1699
 QY 188 ----- 188
 Db 1700 TTCCTGCAAGCTTTCAGGGGTGTTCTTATGACCCAGGGGCTCCACGAACCCAGGCTCTCC 1759
 QY 188 ----- 188
 Db 1760 ACCCCAGGCTCAGCTGAGCAGACACACTGACCTCCCCAGCTAAAGGCGCTGTGGGCTC 1819
 QY 188 ----- 188
 Db 1820 CCTGTCCGACCTGTGGGCGCAATGGCCCTCCCTACTGTAGAGTCCGGTCTCATACCT 1879
 QY 188 ----- 188
 Db 1880 GACCCGTAATGAGAGTCTGTGTGTGCTGCTGCTGCCCACTAGGGCGGCACCCAGCCCT 1939
 QY 188 ----- 188
 Db 1940 GGGCTAAAGCTGGTGTGTGTGTGTACCCCGGGGAGCCCTCCCGACGTGAGAGCCG 1999
 QY 189 ----- 189
 Db 2000 GGTCCCTCCCTCCCAACCCCTGACAGCTACAGGCTGTGCTCCCTCCCAACCCCTGCGGAC 2059
 QY 201 aLlIeGlnGlyLnaSnpheValThrArgArgArgLleaArgArgThrValAlaLeuAlaL 221
 |||||
 Db 2060 TGTATCCAGGGCCCTGAGCTCTGTGACCGGAGGCTATTCGGAGACGGTGGCCAGAGCCC 2119
 QY 221 euleuProGlyArgLeuProGlyArgPro-TyrAlaLeuMetAlaLysTyrIleLeu 240
 |||||
 Db 2120 TGCC-GCGCGTGGCCCTGACAGGCGGACCGGACCTGCTCATGTGCGCAAGTACATCATG 2178
 QY 241 AspLeuGluArgLeuHisProAlaAlaThrThrPheArgValGlyLeuProGly 260
 |||||
 Db 2179 GACCTGAGCGGCTGTGATCCACCGGGGCGCGGACCTTCCACGTGGGCTCCCTGGG 2238
 QY 261 Ala-----GlnGluGluProGlyLeuLeuArgValAlaGlyAspAsnGlyIle-Pr 277
 |||||
 Db 2239 GCCCTGTGTGGCCACGACGCGGCTGGGGCTTCCGCTGTGCTGTGACGGCGCATCCGC 2298
 QY 277 0----- 277
 Db 2299 TGGACCCAGGAGAAACAGGGGTGAGCAGGCTTCCCCCTGCGGGGCCAAATGGGGGGG 2358
 QY 277 ----- 277
 Db 2359 GTTTGGGAGGGGCGGAGATTGTAGGGGATGTGGGGCGGAGCCAAACGAAGACTTG 2418
 QY 277 ----- 277
 Db 2419 GGGAACTGGCGAGGCTTAATGAGGGCGGGGCTTAGTGAAGGAGAGACTGCGGGAATG 2478
 QY 277 ----- 277
 Db 2479 GGAGGGCAAACTGAGTGAAGGGTTGGGCTGAGCGACCCGGGAATGAATAATGGAGGGGT 2558
 QY 277 ----- 277

Db 2539 TGAGAGTGGGTGGTGGCGCTGGAAGCAGCAAGCTAAGAGCAGAGTGAAGCTCCG 2598
 QY 277 ----- 277
 Db 2599 AGAGCTGGGAGGCTGGCAACCGGGGCGAGGATCAAGGGGGGCTCCGCGCAGAGAAAG 2658
 QY 278 -----TyrSerSerAsnAspLysPhe----- 285
 |||||
 Db 2659 AGCGGGGATTTGGCAGAGCGGGGTGGAGATGTTTGGGGATGGGGTAGGGAGGTGTGAGA 2718
 QY 285 ----- 285
 Db 2719 GGTGGAGTTGATAGAGGGGCGGGCTTGGAGGGGTTGAATGCAAAAGGATAGGAG 2778
 QY 285 ----- 285
 Db 2779 TGGATGTGTGGCTTGGGGGTGGTTCATGGGCGGTGTTTGGCGGACCCGCTGGGCC 2838
 QY 286 -----GlnThrPheCysAspPheProGluIleVal 295
 |||||
 Db 2839 CACTAGGTAANCCCAACACCCAGCTCTCCAGCCCTTCTGCGACTTCCAGAAATCGT 2898
 QY 295 LAspValSerIleAsnGlnAlaProArgValGlyProAlaGlyGlnHisArgLeuValTh 315
 |||||
 Db 2899 AGACATTTAGCATCAGCAGAGCCCGCGCTTGGCCCGGAGAGCACCCGCTGTCTAC 2958
 QY 315 rValThrArgMetAspGlyHisIleLeu----- 324
 |||||
 Db 2959 TGTATCCAGAGACAGAACCAAGATTTAGTGTGTGTCAGAGATTCCCTCCCTCAGCCT 3018
 QY 324 ----- 324
 Db 3019 TACCCCGAAGGGCGGACCGGACCCCTCGGGTTTCACTGGGCTGACGCTTGTCCCTCG 3078
 QY 325 --GluAlaGluPheProGlyLeuProGluAlaLeuSerPheValAlaLeuValAspGlyT 344
 |||||
 Db 3079 CAGAGCGCAGATTCACAGGGCTGCGCGAGGCTGTGCTGTGGCGCTGTGACAGGCT 3138
 QY 344 rPheArgLeuIleCysAspSerArgHisTyrPheCysLysGluValAlaProProArgL 364
 |||||
 Db 3139 ACTTCGGCTGACCAACGAGGCTCCAGCACTTCTCTGTGCAAGAGGTGGCACCGCGGAGC 3198
 QY 364 euleuGluGluAlaAlaAspValCysHisGlyProIleThrLeuAspPheAlaIleHisL 384
 |||||
 Db 3199 TGCTGGAGAGAGTGGCGGAGCATGTCAGGSCCCCATCTCTGAGCTTGGCATYCAACA 3258
 QY 384 yLsLeuLysAlaAlaGlySerLeuProGlyThrTyrIleLeuArgArgSerProGlnAspT 404
 |||||
 Db 3259 AGCTCAAGACATGGGGGCTCACGCTCTGCTCTATGTTTCCGCGCAGCCCGCAGACT 3318
 QY 404 rAspSerPheLeuLeuThrAlaCysValGln-Thr----- 415
 |||||
 Db 3319 TTGACAGCTTCTCTCTACTGTCTGTGTGTCAGGTGCGTACTGCTAGGGTGGGTACTGG 3378
 QY 415 ----- 415
 Db 3379 AGGCTGCTGGACAGAGGTTTACCTCTGAATTGAGATTACTCAGAACTACTCAGCAT 3438
 QY 415 ----- 415
 Db 3439 TTGGTTCTCGAACAATAATGAGGAAACGACACCCCGGTTAAGGGGAACAGCGTATTCT 3498
 QY 415 ----- 415
 Db 3499 CTGGCTTAATFACCTGAAATTAGTTCTGTTGCTTTCAGAGATAAAGACTGAACCTGGGCGC 3558
 QY 415 ----- 415
 Db 3559 TGTGCGGCTCCCTGTTATCCAGCCCTTTGGAGGCTGACGTAAAGCATATGCCGAGATC 3618
 QY 415 ----- 415

Db 3619 GGGATATCAACCCATCTGTCGCAATGTTAAACCCCTGTCACCTGTAATATGTAA 3678
QY 415 ----- 415
Db 3679 CTCAGCTGGCAGTGTATGGCATCTGTATATCCACCTACTGCGAAGCTCAAGCAT 3738
QY 415 ----- 415
Db 3739 GTTTTCTAAGTTTAAATAAGGGGGGGGTTTGGGTTTAAAGGAAAAA 3798
QY 415 ----- 415
Db 3799 AAGGGGTGGGAAAAAAGGGGTAAACCTTTTATAGTCTGAGAGCTCTGGCAAC 3858
QY 415 ----- 415
Db 3859 CTGCTGGATGGGGGCTGCACGTAGATGGGTGGAGTCACTTCCTGCTGTAC 3918
QY 415 ----- 415
Db 3919 CCAAAATTTGGGGGGTGGGGTGGCCAGTTAAAGTTAAGGTTGNAACCTTCGGG 3978
QY 415 ----- 415
Db 3979 GCCCAAAAGGGGAACCCCTTTTAAATAATTTTGGGGGGTAAG 4038
QY 415 ----- 415
Db 4039 GGGATTTTTTCCCCCAATGGGGGGGGCGGGCANTTTTAAACCCGGGG 4098
QY 415 ----- 415
Db 4099 GGGGGGGTCCAGGGGCCAGGATAACCTTTAATTTCCCCCGGGAAAGGCCA 4158
QY 415 ----- 415
Db 4159 AATCTTTTCCCAAAAAAAAAAAGAAAGAGAGACTGAGAGAG 4218
QY 416 ----- 416
Db 4219 AGTGTCTGCTGCTACGTCCTCAGGGGCCACTCTTTCAGACACCCCTTGCTCT 4278
QY 420 AspyrlysglycylleuilearglnsproseryalaPhaserleuValglyleu 439
Db 4279 GATTATAAGGCTGCTCATCCGCGCACGCCACAGAACCTCTCTGTTGGCTC 4338
QY 440 SerGlnProHisArgSerleuArgIuleuAlaIaCysTrpAsnSerIlyleuArg 459
Db 4339 AGCCGACCCCAAGCAGCTTCTGAGAGCTCTGCAACCTGCTGGGATGGGGGTGCAC 4398
QY 460 ValAspGlyAlaLeuIleuThrSerCysAlaProArg-Prolys- 476
Db 4399 GTAGATGGGGTGGCAGTCACTCTGCTGTATCCCAAGCCAAAGGTGAGGCC 4458
QY 476 ----- 476
Db 4459 CCTGTGCCCTGGGAATGAGTGGCTGATCTGGACCCCTGTTCTATGTGATACAG 4518
QY 476 ----- 476
Db 4519 CTCCTGTGTGATGGCAAGTGTAGAACTGCAGATCATGTGGTTAGGAAGAAAGG 4578
QY 476 ----- 476
Db 4579 CGATTTGTGCTTCAGGAAGTTAGACCCTAACCCTAAGGCTGATTTGCCACGGTA 4638
QY 476 ----- 476
Db 4639 CAGCTAGATGTAGCCATGTCACCTGACCTTGNTTATCTCAGAATCTCACTTCTTC 4698
QY 476 ----- 476
Db 4699 TTCTTGGGTGACTTCACTCTTGGCAACAACCTCCCTGTATGGGACAATTAATTCT 4758

QY 476 ----- 476
Db 4759 GAGTATGTGAGACCCCTTCTGAGTCTTCCAGCAATTCCTGGTAAGAACCCGAT 4818
QY 476 ----- 476
Db 4819 TCCTCGGCTGCTCCAAAGTGTCTAATTTGGCTTCAATTAATAGGCCCTGA 4878
QY 476 ----- 476
Db 4879 AAGGGGTTTCCTTCCCTTCCCTGGGGGTTTCACCCCAACAAATTCCTCATG 4938
QY 476 ----- 476
Db 4939 GGGGGGTCTTCTTGTGATTTTCCCTTCCCAAAATTAACCATGTATAGTTGCG 4998
QY 476 ----- 476
Db 4999 GCCACAGAAAGATGGAAGTGGAGAGAGAGAGATGCGGGGTTCTTCCACTCC 5058
QY 476 ----- 476
Db 5059 AATATCCAAATTTGTGTTCACAGACCTGCCTTAGGGSAAACCTGGGCACAGTCA 5118
QY 476 ----- 476
Db 5119 CTTGGCCCTGAGATTGCGATTAGTGCACACAGCANTACATGATTTTGTATAGG 5178
QY 476 ----- 476
Db 5179 ACAGAGTCCCACTATCATATGCTTTTGTGTTGATTGCTTGTGTTGAGACAG 5238
QY 476 ----- 476
Db 5239 AGTCTCTTCTGTCACCCAGGCTGAGTGCAGCGCGCATTTTGTATCTCAACTTC 5298
QY 476 ----- 476
Db 5299 CGCTCCCAATTTAGACATTTCTCTGCTTCAAGCTCCCGAGTACTGGATTAACAG 5358
QY 476 ----- 476
Db 5359 GCATGCACACACGCGCGCTAATTTTGTATTTTAGTAGAGAGGGTTCCTCATG 5418
QY 476 ----- 476
Db 5419 TTGTCAGCTGTGTCTGACTCCCGACCTAGATGATCTGCCGCTCAGCTCCCAA 5478
QY 476 ----- 476
Db 5479 GTGCTGGATTACAGGGCTGAGCCACGCGCGCTAATTTTGTGTCAAGACAG 5538
QY 476 ----- 476
Db 5539 GTCTGCTATGTGCTAGGCTGTTTGAACCTGTGCTCAAACCATCTCTGCTCA 5598
QY 476 ----- 476
Db 5599 GCTCCCAAGTGTGAGACAGGTTTCCCGTTCTTTAAACCAACCAACCCCG 5658
QY 476 ----- 476
Db 5659 CCCCAATGTGTGACCTTTTATTAACCCCAACTTTTGGGGGGGAGGGGGGTT 5718
QY 476 ----- 476
Db 5719 TAGTCCCCCGGGTCTTTTGGGTTTGGCAACCCCGCCCTTAAAAAAGGGGCA 5778
QY 476 ----- 476
Db 5779 ACCCTTTTCCGCTAATAAAGAAAAAGAGTGTGGAGGTTGAATAATTCGT 5838

```
QY 476 ----- 476
Db 5839 TTTTCCCATAAATCCGTGGGAGAGGCGGGGGGCTTACTTGAACCCCGGGGCTCCGGT 5898
QY 476 ----- 476
Db 5899 GTGGTTCCTCCGATAAACACAGAAATTTCCCGGGGAAACCGGGGAAATTTT 5958
QY 476 ----- 476
Db 5959 TTTTAAAGAACAAAAAATTAGCGGGGACGGGTGGACACACCTGTATGCCAGC 6018
QY 476 ----- 476
Db 6019 TATTGGGAGCGGAGGAGGAAATCGCTTGAACCTGGAGGTGGAGTTGCAGTGAGC 6078
QY 476 ----- 476
Db 6079 TGAGATCGACCATTTGCCACCCAGCCTTGATGACAGAGTGAGATCCGCTTCAGCAGCAG 6138
QY 476 ----- 476
Db 6139 CAGCAACAGCAAAACAAACAAACAAAGCATGTGCCCTGAAGTCTTCATCTCAG 6198
QY 476 ----- 476
Db 6199 GGTGGGCTTAGAGGGTACTCAACTAAGCATGAGTAGCTAACCTTGGGGACTTT 6258
QY 477 ----- 476
Db 6259 TCACCTCTGATTTCTGGTTTTTTCCTCCCTACCTCTCTCCCATAGAAAAGTCCACACGAT 6318
QY 482 eValValArgArgLycysnProAlaProAlaProGlyCysSerPro---SerCysCy 501
Db 6319 CGTGCTCAGAGAGGTACACAGCCACCATCATCTCTTGCTTACAGCCCAATCCCAATA 6378
QY 501 salaleuThrlInleuSerPheHsThrIleProThrAspSerleuGluIyr----- 518
Db 6379 CCACCTGATCAGATGACATTTTCCAAAGATCCCTGCTGACAGCCTGGAGTGGTAAGAG 6438
QY 518 ----- 518
Db 6439 CCCTGGAAATGAGGCGATACCTAGTCTGGGCTCCAGAGACTCAGATGCTGCCTCAG 6498
QY 518 ----- 518
Db 6499 GCATATGCTAATTTTACCTTGCCTTCGGTTTCCCATCTGTAAATGGGGCCAGCAGCT 6558
QY 518 ----- 518
Db 6559 ATGTCTCGCTTGGGCTGGGATCCTGCAGAACCCCTCACTGGCCTTCTGTGCTGCC 6618
QY 519 ----- 518
Db 6619 TCACCATTCAGCATGAGAACTGGGCGCATGGGCTTCTCACCAAGATTTTACCGGGCTGTC 6678
QY 535 rGArGluValAlaAspGlyIuThrHsAspSerGluValLeuLeuLysValIleTAspS 555
Db 6679 GCCATAGCTGGTGGATGGGAGGCGGCAAGACAGAAAGTGTGCTGAAGGTCATGATG 6738
QY 555 eFAlHsIArgAsnCyMetGluSer----- 563
Db 6739 CCAAGCACAGAACTGCATGAGGT- GAGAGGCAATGTGACACAGACTTTGGAGTGGG 6797
QY 563 ----- 563
Db 6798 GCGTGGCTTGACCGGGGCTCGTGATGCAAGACATTTACAACACACAGGAGCCTGGCG 6857
QY 563 ----- 563
Db 6858 GTGGTGGCTGATGCTGCATCATAGCGGGAGCTCGTCTTGACAAAAAATTTTAGGCC 6917
QY 563 ----- 563

Db 6918 GGGCGCGGTGCTCACGCTGTATCCAGCAGCTTTGGGAAGCTGAAGCGGAGAAATCA 6977
QY 563 ----- 563
Db 6978 TCGAGTGGGAGTCTCGAGACCAACCTGCCCAACATGAGAGAAACCGTCTCTATAA 7037
QY 563 ----- 563
Db 7038 AAAACCAAAATTTACCGGGCGTGGTGTGCATGCTGTATATCCAGCTACTCGGGAGTTC 7097
QY 563 ----- 563
Db 7098 TGAGGACAGGAATCTCTGAATCTGGAAAGGAGGTTCCAGTGATCAGATCGCGCC 7157
QY 563 ----- 563
Db 7158 ACTGCACTCAGCCTGGGTGACAGAGGAGACTGTCTCAAAAAACAANCAANCAANA 7217
QY 563 ----- 563
Db 7218 CAAAAACCATAGATGATAGTGGAACTTGTGCCGTATCAGAAATCATGTAGTGTCT 7277
QY 563 ----- 563
Db 7278 GTGTGCACTAATGACAGACTCCAGGGCCAAAGTGACCTGTGGCCAGTTCCTTAAG 7337
QY 563 ----- 563
Db 7338 GCAGGTCTGTGAGCAAAAATTTGGGATTTTGGAGTGGAAAGAACCCAGCAGCTTCTC 7397
QY 564 ----- 564
Db 7398 TCCCTTCCACCTTCCAGTATCTCGGAAGCGCAGCCTTGATGAGCAAGTCTGT 7457
QY 576 yPProHsLeuValLeuLeuHsGlyValLysMetAlaGlyAsp----- 590
Db 7458 ACCGCGATCTGTGTCTGCCAGCGCGTGTGCATGTGCTGGAGACAGGTGACTGCTCTAG 7517
QY 590 ----- 590
Db 7518 GGGGTGTTAAAGGGAGAGGAGGAGAAAGAGGCGAGGGGAGAAAGAGCTAGGGGCA 7577
QY 590 ----- 590
Db 7578 GAGAGAGAGAGAAAGCGGGGAGAGGCGAAACGGAGACAAAGAGAGGTTAGCGAGA 7637
QY 590 ----- 590
Db 7638 GGAATAAAACTGCGCTGACGTTTATAAAGATAGGGGGAGAGAGCTGCTCGAGGC 7697
QY 590 ----- 590
Db 7698 AAATTATATAGGCCCGGAGAGAGACTTGTGGCTTGGGGGGGCTGAGAGAGAGACCC 7757
QY 590 ----- 590
Db 7758 GCCCTTTAAACAAACTAGAAAGACCTGGGGATATATCCCGCTTATTTATTAACCTGG 7817
QY 590 ----- 590
Db 7818 AGAGAACTCAAGAAAAAGAGATCCCTGGGATGCCCCAGATATTTTCGGAGTGGCG 7877
QY 591 ----- 591
Db 7878 AAAGAACTAAACTTGGAAATCTGAGCAATATGGCTGAGACACATGGTGGAGAAATTTGG 7937
QY 598 aLtyrIleuGlyAlaIleAspMetTyrLeuArgLysArgLysIleuValSerAlaSerT 618
Db 7938 TACACCTGGGGGCGATGACATGTATCTCGAAACGTGGCCACCTGGTGGCCAGGACACT 7997
QY 618 rPlysLeuGlnValThrLysGlnLeuAlaTyrAlaLeuAsnTyr----- 632
|||||
```

Db 7998 GGAAGCTGCAGGTGTCAACACAGCTGGCCTACGCCCTCACTATCTGTAGTCTCTC 8057
QY 632 ----- 632
Db 8058 CCCCAGGAGATTGTGTGGGGAGGGGGGTTAAAAAACCCAAACCCCTTAAT 8117
QY 632 ----- 632
Db 8118 TTTGGGGGGGGTGTGGGTGTTCCTCCCTCATTCCTCCGGGAAGTTTTTTTTTT 8177
QY 632 ----- 632
Db 8178 AAAGGGAAGAAGGATGTTTTTTTTTTTAAAGATTGTGAAGTTGGGGGGGG 8237
QY 632 ----- 632
Db 8238 CATTTTACTTTCTTTTTGGAAAAATGAAAAAAATTTTTTTTTTTTNGGGT 8297
QY 632 ----- 632
Db 8298 TTGGGGAAGGGGGTGGCCGGGAAGAGTTTCTTTTGGAGAAAAAGGACATGGCC 8357
QY 632 ----- 632
Db 8358 CTTACCCCTGAGTGGGTTTTTTCATGAAGAAGGAGAGGATTTGGGGAG 8417
QY 632 ----- 632
Db 8418 ACCAAAGCAAGCCGGCGGGAGAAAGGAGATTTCCCTGGCCCTATCAGCATT 8477
QY 632 ----- 632
Db 8478 TCCGTTGCCAGAGACAGAGGAAAAAGTGGAGAGGAATTAATTGGAAGCTTGAGCAG 8537
QY 632 ----- 632
Db 8538 GGGCTGGAATCCAGGTAGAGGTTAATTTTTTCAAGGATTTGGGTTAACAG 8597
QY 632 ----- 632
Db 8598 ATGGGTGTGTGTGTAGAGGGGAGACATACATGTCATTTTGGGGGTGAGTAGAG 8657
QY 632 ----- 632
Db 8658 GGGCAAGACAGATATCAGATTTTCAAGGACATGTGTGAGCATTTTGGAGCAGTACAA 8717
QY 632 ----- 632
Db 8718 GTGGGTTTGAAGATATAGAGTTTGGCAACAGACTCTTCATTTATCAACCTCC 8777
QY 633 ----- 633
Db 8778 CCGGCATTTCGTGTGTGTGGCCCCCTTAGAGAGACAAAGCCCTGCCCATGTGC 8837
QY 644 SerAlaArgValLeuLeuAlaArgLugLylValSpLysnProPheIleLys 653
Db 8838 TCTCCCGGAAGTGTCTCGGCTCGGAGGAGGCTGATGGAGCCGCCCTTCATCAAG 8897
QY 664 LeuSerAspProGlyValSerProThrValLeuSerLeuGlu 677
Db 8898 CTGAGTACCTCGGGTACGCCCGCTGTGTAAAGCTGAGGTAAAGTCTCTGAGAGTGG 8957
QY 677 ----- 677
Db 8958 AGAAGGAGGGGCTGAACAGGCAAGAGATGATCCCTGATCCACTTCATTCCTCA 9017
QY 678 -MetLeuThrAspArgLileProTrpValAlaProGluCysLeuGlnLualagInThrle 697
Db 9018 GATGTCACCGACAGATCCCTGGGTGGCCCGAGAGTCTCCGGAGGAGCGAGACT 9077
QY 697 ucysLeuGlnLualaspyLysTrpGlyPheGlyAlaThrThrTrpGluValPheGlnArgL 717
Db 9078 TAGCTTGAAGCTGACAAGTGGGCTTCGGCCACGGCTTGGGAAGTGTTAAGTGGCGT 9137

QY 717 yProAlaHisIleThrSerLeuGlnProAlaLys 728
Db 9138 CACATGCCCATCAGTACGCTGTGATGCTAAGGTGACAGCCCTCACCGGCATGGT 9197
QY 728 ----- 728
Db 9198 CTCCGAACCCCACTTTGACAGCCGGGCACTGACATCCACTGTGGAGATTGGGGTG 9257
QY 729 ----- 729
Db 9258 GGTCTATTGGGTTGGGATTACGACTGCTCTCTACCTCAGAAACTCCATTTTATG 9317
QY 734 LyspGlnGlyLeuLeuProAlaLeuLysTrpThrGluLeuAlaGlyLeuIleThrGlnC 754
Db 9318 AGAAGCGGACAGACTGCTGCCGCCCAAGTGGACAGAGCTGGCCCTGCTGATTCAACAGT 9377
QY 754 yMetaLarYraspProGlyArgArgProSerPheArgAlaIleLeuArgAspLeuAsnG 774
Db 9378 GCATGGCTTATGAGCCGGTCCAGAGGCCCTCTCCGAGCCGTCATTCGTACCTCAATA 9437
QY 774 LysLeuIleThrSer 778
Db 9438 GCTCATCTCTCAGGTGCCCTGGAGCGGTTGGGTGGGAGGGCTGTGATGTATAT 9497
QY 778 ----- 778
Db 9498 TGGCCCAGTGAAGAGCGTGTGTCAGTAGGCCACAGCCCTGTGTGTGTAGTT 9557
QY 778 ----- 778
Db 9558 GGAGGGTTGATGACTATGACAGACAGTGTAGCTTCAACAGTACTGTAAGCAACAGTC 9617
QY 778 ----- 778
Db 9618 CCTCAGTTCAAGGTTCACATGATGATTTATTTATTTATTTATTTTGAACAGAG 9677
QY 778 ----- 778
Db 9678 TCTTGTCTGTGGCTAGAGCTGACGTGACAGTGCACAAATTTGCGTCACTGAACTCTA 9737
QY 778 ----- 778
Db 9738 CCTCCTTAACCTGATGTTCAAGGATTTCTCTCCTCAGACTCTGATAGTGGATT 9797
QY 778 ----- 778
Db 9798 ACAGGTGTGCAACAGCCAGCTAATTTTTTTTGGTATTTTGGAAAGTTTGGGCT 9857
QY 778 ----- 778
Db 9858 TTCAACCATATTTGGCTAGGGTGTCTTGAACCTTGACCCCAAGGTGATCCGCCAC 9917
QY 778 ----- 778
Db 9918 CTCTACCTCTAAGAGTGTGGATTCAAAATTTCTCTACTGTGCTTGTGTGATAT 9977
QY 778 ----- 778
Db 9978 TGCCAATTAGTCGATATAGGTTTGCACAGTAGACTCATGATTTCACTTCGTGAG 10037
QY 778 ----- 778
Db 10038 ACTTTGTAACAAGAGTTTCAAGTGTCAAGAGGAGACCGGATTTGTTATAAATAA 10097
QY 778 ----- 778
Db 10098 CCTGCAGTATTAATAGAAATCAATAAATAAATAGACTCAGTATATACAGGAGAACAC 10157
QY 778 ----- 778
Db 10158 CCAATTAAGTTTATGTTGCCCTTAAGAAAGAAAAATGGCGTGAATGTAATAATATTA 10217

polyA_site		/note="17 a nucleotides"			
BASE COUNT	1554	a	853	c	1016
			g		1389
			t		
Alignment Scores:					
Pred. No.:	1..1e-152				4812
Score:	2637..50				535
Percent Similarity:	66..15%				194
Best Local Similarity:	48..55%				343
Query Match:	45..01%				30
DB:	4				14
US-09-397-967-16 (1-1099) x AB006011 (1-4812)					
Oy	16	CysSerLeuSerSerSerGluAlaGluAlaLeuValLeuLeuPro-ProAlaGlyPr	35		
Db	409	TGTGAAGCAATATATATCCAGTCCACAGGCTATCTTTATCACTCCCTGGAGAAAGTGA	468		
Oy	35	oGlyProGlnArgLeuSerPheSerPheGlyAspArgValLeuAlaGluAspLeuGly	55		
Db	469	GGGC-----GATTATTTGAAGTTTCCAGTGGAGAGATGTGGCAGAGAAATTTGGT	522		
Oy	55	IArgAlaAlaGlyAspGlyLeuLeuProValTyrHisSerLeuPheAlaLeuAlaIh	75		
Db	523	TCTGCTCTTAAGTGTGTGGATCACACAGCTGTATACATAGATATTTGCTTTAATGAA	582		
Oy	75	rGluAspPheSerGlyTrrPheProGoserHisIlePheCysIleGluAspValAspIh	95		
Db	583	TGAACACAGAAAGATCTGTATCCACCCACACATGCTTCCAGTGAATGATCAATCTACAG	642		
Oy	95	rGlnAlaLeuValTyrArgLeuArgPheTyrPheProAspTrpPhe-----GlyLeuG	113		
Db	643	GCCAAATGATCTACAGATACAGATACGATTTTACTTTCTTACGTGATCTGAATGACAGAA	702		
Oy	113	uTrpCysHisArgPheGlyLeuArgValAspIleThrSerAlaIleLeuAspLeuHisIya	133		
Db	703	CAGAACCTATCCGATCGATGAAATCTTCAGAGAGTGTGAACCTCTCTGTATGACTTTGT	762		
Oy	133	IleuGluHisIleuPheAlaGlnHisArgSerAspLeuValSerIleArgLeuProValG	153		
Db	763	CATGTTCTTACTTTTGTCTCAGTGGCAGATATTTTGTGTATGATGATGGTAAAGATACC	822		
Oy	153	yluSerMetLeuGluGlnGlyArgPheLeuSerLeuAlaValLeuAspLeuAlaGlnIe	173		
Db	823	TGTATACATGAACCTCAGAGAAATGCTTGTGGATGGCAGAGTGAATGATGAGAGAAAT	882		
Oy	173	IAlaArgGlnAlaGlnArgProGlyIleuLeuLeuGlyThrValSerTyrGlyAlaGly	193		
Db	883	AGCCAAAGAAAGATCAACCCCACTGACATCTATTAAGTCTTGTGCACCTCAAGACATT	942		
Oy	193	sLeuProProSerLeuArgAspValIleGlnGlyGlnAspPheValIhArgArgArgI	213		
Db	943	CTTACCAAAATGTGTGTGACCAAAAGATCCAAAGCTATCATATTTTGACAGAAAGCAAT	1002		
Oy	213	eArgArgThrValValLeuAlaLeuLeuProCysGlyArgLeuProGlyArgProTyrAl	233		
Db	1003	AAGGTACGATTTTGGCAGATTATTTAGTGAATTCACGCATCTGCAGAAACCATGCCAGAA	1062		

Oy 233 aleuMetAlaIySTyriIleuAspIeuGluArgIeuHisProAlaIaIeThrThrGluTh 253
 Db 1063 GTTGAACCTTAATATCTTAATCTGGAACCTTCTGCTGCTCTTCTACACAGACA 1122
 Oy 253 rPheArgValIglYLeuProGlyIle-----AlaGInGluGluProGlyLeuLeuArg 269
 Db 1123 GTTTGAAGTAAAGAACCTGGAAGAGGTCCTCAGGTGAGAGATTTTTCGAACCTATT 1182
 Oy 269 gValaIaIglYAspArgnGlyIleProTrpSerSer----- 280
 Db 1183 AATAACTGGGAACGCTGGAATTCAGTGTCAAGAGGGAACATAAAGAAAGTGAGACACT 1242
 Oy 281 -AsnAspGluLeuPheGlnThrPheCysAspPheProGluIleValAspValSerIleAs 300
 Db 1243 GACAGAACGAGATTACATTATATTGATGATTTTCTGATTCATTGATGATGATGATTTAA 1302
 Oy 300 nGlnAlaProArgValIglYProAlaGlyIleHisArgLeuValThrValThrArgMetAs 320
 Db 1303 ACAGGCAACCAAGAAAGGC---TCAATATGAAGACAGATTTGATACATTATACAGACAGA 1359
 Oy 320 pGlyHisIleLeuGluAlaGluPheProGlyLeuProGluAlaLeuSerPheValAlaLe 340
 Db 1360 TGGCAAAAGCTGGAAATTTGAACCTTAGCTCATTAAGAGAGAGCTTCTTTGTATCGTT 1419
 Oy 340 uValaIaSPGlyTrPheArgLeuIleCysAspSerArgHisIleTyrrPheCysLeuGluValAl 360
 Db 1420 AATTGATGGATATTAGATTATACCTGACATGACATCTTACCTTTGTAAACAAATAC 1479
 Oy 360 aProProArgLeuGluGluGluAlaSPValCysHisIglYProIleThrLeuAspPh 380
 Db 1480 ACCTCCAACTGGCTTGAATAATATACAAAGCAACTGTCTAGGCCCAATTTCAATGATTTT 1539
 Oy 380 eAlaIleHisIleGlyLeuAlaIaIaGlySerLeuProGlyThrThrIleLeuArgIle 400
 Db 1540 TGGCAATCAGTAACAGAAAGACAGATATACAGACTGAGCTTGTGTATCTGATGACAG 1599
 Oy 400 rProGlnAlaPtyrAspSerPheLeuLeuThrAlaCysValGlnThrProLeuGlyProAs 420
 Db 1600 TCTTAAGACCTTAATTAATATTCTGACTTTTGGTGGTGGACGGAAGAAATGTCACTGA 1659
 Oy 420 pTyrrIleGlyCysLeuIleArgIleAspProSerGlyAlaIaIaPheSerLeuValGlyLeu 440
 Db 1660 ATTTAAGCACTGTTTATTCACAAAATGAGAAATGCGAATACACCTGAGTGTACTAG 1719
 Oy 440 rGlnProHisArgSerLeuArgIleLeuAlaIaIaCysTrpAsnSerGlyLeuArgVa 460
 Db 1720 GAAGAACTTCACTTAACCTTAAGATCTTTGAAATGTTTACCAAAATGGAACGTGTGCTTC 1779
 Oy 460 IaSPGlyAlaAlaIleuTyrrLeuThrSerCysAlaProArgProGlyGluIleYSerAs 480
 Db 1780 AGACAGATTAATTTTCCAGTTTACTTAATGCTGTCCCCCAAGCAAAAGATTAAGTCAAA 1839
 Oy 480 nLeuIleValValArg---ArgGlyCysAsnProAlaProAlaProGlyCysSerProse 499
 Db 1840 CCTTCTAGTCTTGAACCAATGATTTCTGATGTCGCCAACCTTCACCAACATTACAGAG 1899
 Oy 499 rCysCysAlaLeuThrGlnLeuSerPheHisIleProThrAspSerLeuGluTrpPh 519
 Db 1900 GCATTAATATGTGACCAAAATGTGTTTCCAAATATCAGAAATGAGATTTGATATTAA 1959
 Oy 519 sGluAsnLeuGlyHisIleGlySerPheThrIlePheArgIleSerIleArgIleValVa 539
 Db 1960 TGAAGCTTGGCCCAAGGCAAGCTTTACAAAATTTTAAAGGTGTAAAGAGAGAGTAGG 2019
 Oy 539 IaSP---GlyGluThrHisAspSerGluValLeuLeuValMetAspSerArgHisAr 558
 Db 2020 AGACTAGGTCAACTGATGAACAGAGAGTCTTTTAAAGTGTGATTAAGACACATAG 2079
 Oy 558 gAsnCysMetGluSerPheLeuGluAlaIaIaSerLeuMetSerGlnValSerTyrrProH 578
 Db 2080 AACTACTAGAGCTTTCTTTTGAAGCAGAGCATGATGAGCGAGCTTCTCTATTAACA 2139

Oy 578 sLeuValLeuLeuHisGlyValCysMetAlaGly---AspSerIleMetValGlnGluPh 597
 Db 2140 TCTGCTTTTAATTAATTAATGAGATTAATGCTGTGTGGAAGAGAAATATTCTGCTCAGAGAT 2199
 Oy 597 eValTyrrLeuGluAlaIaIaLeuAspMetCysLeuArgIleGlyHisLeuValSerAlaSe 617
 Db 2200 TGTAAATTTGGTCACTGATGATATCTGTAAGAAAGATTAATAATTAATTAATATT 2259
 Oy 617 rTrpIleLeuGlnValThrIleGlnLeuAlaIaIaLeuAsnIleTyrrLeuGluAspIleG 637
 Db 2260 ATGGAACCTTGAAGTGGCTTAACCTTGGATGGCCATGCGCATTTTCTAGAAGAGAAAC 2319
 Oy 637 yLeuProHisIglYAsnValSerAlaArgIleValLeuAlaArgIleGlyIle----- 655
 Db 2320 CCTTATTCATGGGATGTGTGGCCCAAAATTTCTTCTTATCAGAGAAAGACAGAGA 2379
 Oy 655 -AspGlyAsnProProPheIleIleLeuSerAspProGlyValSerProThrValLeuSe 675
 Db 2380 GACAGGAATCTCCCTTCATCAAACTTAGTATCTGTCATGATTAATACAGTTTGGCC 2439
 Oy 675 rLeuGluMetLeuThrAspArgIleProTrpValAlaProGlyCysLeuGlnGluAlaG 695
 Db 2440 CAAGGACATTTGCAAGAGAGAAATACATGGGTGCTCTGTAATGATTAAGAAATCTTA 2499
 Oy 695 nThrLeuCysLeuGluAlaAspIleStrpIlePheGlyAlaThrThrTrpGluValaPheG 715
 Db 2500 AATCTAACTTGGCAACAGATTAATGAGATTGTGTACCTTTGTGGAAATCTGCAG 2559
 Oy 715 nArgGlyProAlaHisIleThrSerLeuGluProAlaIleGlyIleGlyIleGlnCysWe 735
 Db 2560 TGGAGGTGATTAAGCCCTTAATGCTTTGGATTCTCAAAAGCAATACATTTTATGAGAA 2619
 Oy 735 pGlnGlyIleLeuProAlaLeuIleStrpThrGluLeuAlaGlyIleThrGlnCysWe 755
 Db 2620 TAGGCACTACGCTTCCGCAACCAAGTGCACAGATTAAGCAAAATTTATTAATTTGAT 2679
 Oy 755 tAlaTyrrAspProGlyArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLe 775
 Db 2680 GGAATTGAACAGATTTTAGACTTCTTTCAGAGCCATCATACGAGCTTAATAGCTT 2739
 Oy 775 uIleThrSerAspTyrrGluLeuLeuSerAspProThrProGlyIleProSerProArg-- 794
 Db 2740 GTTACTCCAGATTATGAACCTACTAACAGAA--AATGATATGTTACCAAAATGAGCAAT 2796
 Oy 795 -AspGluLeuCysValAlaGlyAlaGlnLeuTyrrAlaCysGlnAspProAlaIlePheG 814
 Db 2797 AGCTGCCCTAGAGCTTTCTGCTGCA-----TTTGAAGACCGAGACCTTACACAGTTTGA 2850
 Oy 814 uGluArgHisLeuIleTyrrIleSerLeuLeuGlyIleGlyAsnPheGlySerValGluLe 834
 Db 2851 AGAGAGACACTTGAATAATTTCTACAGCAACTGGGCAAGGTAATTTTGGGAGTGTGAGAT 2910
 Oy 834 uCysArgTyrrAspProLeuGlyAspAsnThrGlyProLeuValAlaValIleGlnLeuG 854
 Db 2911 GTCCCGGTATGACCCCTCTACAGCAACACTGGGGAGGTGGGTGGTGAAGAAAGCTCCA 2970
 Oy 854 nHisSerValProAspGlnIleArgAspPheGlnArgGluIleGlnIleLeuIleValaLe 874
 Db 2971 GCACAGTACTGAAGACAGACACTCAGAGCTTTGAAGAGAAATTTGAATCTTAATTCCTT 3030
 Oy 874 uHisSerAspPheIleValIleTyrrArgIleValSerTyrrGlyProGlyValArgIleSer 894
 Db 3031 GCAGCATATGACAAATTTTGAAGTACAAAGAGAGTGTCTACAGTGTGCGCGGTATCTP 3090
 Oy 894 uArgLeuValMetGluTyrrLeuProSerGlyCysLeuArgAspLeuLeuIleArgHisAr 914
 Db 3091 AAGATTATTAATGAATATTACCATTAAGAACTTTTACAGATTAATCTTGAAGAACTTA 3150
 Oy 914 g---GlyLeuHisIleThrAspArgLeuLeuPheAlaTrpGlnIleCysIleGlyMetG 933
 Db 3151 AGAAGGATGATGATTAATAAAACTTCTTCACTATACATCTCAGATATGCAAGGTTATGGA 3210
 Oy 933 uTyrrLeuGlyAlaIaIaArgArgCysValHisArgAspLeuAlaIaIaArgAsnIleLeuValG 953

[illegible]

GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 28, 2003, 13:46:19 ; Search time 105.732 Seconds
(without alignments)
4679.258 Million cell updates/sec

Title: US-09-397-967-15
Perfect score: 90
Sequence: 1 AKLPLDKDYVYVREP 17

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame_p2n.model -DEV=xlp
-Q/cgnt1/USPTO.spool/US09397967/rnat_26042003.182312.10714/app.query.fasta.1.1486
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsu62 -TRANS=human4.cel -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pic -MORF=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09397967.eccn1.1.5398.etrnat_26042003.182312.10714 -NCPU=6 -ICPU=3
-NO_XLPHY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
1: gb_ba:*
2: gb_hlg:*
3: gb_in:*
4: gb_ov:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_ph:*
24: em_pl:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	90	100.0	165	6 E08798	E08798 cDNA encodl
2	90	100.0	198	6 HSU08340	U08340 Human clone
3	90	100.0	448	6 AX203080	AX203080 Sequence
4	90	100.0	778	6 AX203087	AX203087 Sequence
5	90	100.0	2914	6 AX203084	AX203084 Sequence
6	90	100.0	3620	9 HSU31601	U31601 Human tyros
7	90	100.0	4064	6 AX203081	AX203081 Sequence
8	90	100.0	4064	9 HSU09607	U09607 Human JAK f
9	90	100.0	13562	9 HSU070065	U70065 Human JAK3
10	90	100.0	21323	9 AFS13860	AFS13860 Homo sapi
11	90	100.0	41006	9 AC007201	AC007201 Homo sapi
12	83	92.2	3528	10 M05JAK3A	L40172 Mus musculu
13	83	92.2	3733	10 M05PYKIN	L32955 Mouse prote
14	83	92.2	4016	10 M05JAK3H	L33768 Mus musculu
15	83	92.2	10326	10 M05JAK3H	U71201 Mus musculu
16	83	92.2	218208	2 AC073700	AC073700 Mus muscu
17	83	92.2	223734	2 AC073750	AC073750 Mus muscu
18	80	88.9	3644	5 AF034576	AF034576 Gallus ga
19	76	84.4	3778	10 RATVAK3	D28508 Rat mRNA fo
20	71	78.9	14907	5 AF090382	AF090382 Tetradon
21	69	76.7	4387	5 DRE5690	AJ005690 Danio rer
22	66	73.3	171	5 XU11724	U11724 Xenopus lae
23	64	71.1	166	5 AB062055	AB062055 Gallus ga
24	64	71.1	214	10 M05PTKC	M33423 Mouse prote
25	64	71.1	1967	5 DRE5691	AJ005691 Danio rer
26	64	71.1	3435	6 ARI29823	ARI29823 Sequence
27	64	71.1	3495	6 AR048200	AR048200 Sequence
28	64	71.1	3495	6 AR068116	AR068116 Sequence
29	64	71.1	3495	6 AR071176	AR071176 Sequence
30	64	71.1	3495	6 ARI77639	ARI77639 Sequence
31	64	71.1	3495	6 162349	162349 Sequence 2
32	64	71.1	3495	6 187981	187981 Sequence 2
33	64	71.1	3500	6 AF001362	AF001362 Homo sapi
34	64	71.1	3629	6 ARI36148	ARI36148 Sequence
35	64	71.1	3629	6 ARI43945	ARI43945 Sequence
36	64	71.1	3629	6 192575	192575 Sequence 8
37	64	71.1	3629	10 M05PTK	L16956 Mouse prote
38	64	71.1	3731	10 RNUI3396	U13396 Rattus norv
39	64	71.1	4042	4 AB036337	AB036337 Sus scro
40	64	71.1	4161	9 AF005216	AF005216 Homo sapi
41	64	71.1	4599	4 AB036336	AB036336 Sus scro
42	64	71.1	4812	4 AB006011	AB006011 Sus scro
43	64	71.1	5117	6 AX354568	AX354568 Sequence
44	64	71.1	5117	6 AF058925	AF058925 Homo sapi
45	64	71.1	85671	2 AC0936324_3	Continuation (4 of

RESULT 1

ALIGNMENTS

E08798 LOCUS E08798 165 bp RNA linear
 DEFINITION cDNA encoding novel human tyrosine kinase.
 ACCESSION E08798
 VERSION E08798.1 GI:2176910
 KEYWORDS JP 1995059569-A/1.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1. (bases 1 to 165)
 AUTHORS Sakano, S.
 TITLE NOVEL TYROSINE KINASE PEPTIDE AND DNA CODING THE SAME
 JOURNAL Patent: JP 1995059569-A 1 07-MAR-1995
 ASAHIT-CHEM-IND CO LTD
 OS Homo sapiens (human)
 PN JP 1995059569-A/1
 PD 07-MAR-1995
 PI 25-AUG-1993 JP 1993210404
 RI SAKANO SEIJI
 PC C12N15/09,C12N9/12;
 CC strandedness: Double;
 CC topology: Linear;
 CC hypothetical: No;
 CC anti-sense: No;
 FH Key Location/Qualifiers
 FT source 1.165
 FT /organism="Homo sapiens"
 FT /cell_type="megakaryoblast"
 FT /cell_line="UT-7"
 FT mat.peptide 1.165
 FT /product="novel human tyrosine kinase".
 FEATURES source Location/Qualifiers
 1.165
 /db_xref="taxon:9606"
 BASE COUNT 36 a 56 c 39 g 34 t
 ORIGIN
 Alignment Scores: 9.04e-09 Length: 165
 Pred. No.: 90.00 Matches: 17
 Score: 100.00% Conservative: 0
 Percent Similarity: 100.00% Mismatches: 0
 Best Local Similarity: 100.00% Indels: 0
 Query Match: 100.00% Gaps: 0
 DB: 6
 US-09-397-967-15 (1-17) x E08798 (1-165)
 Oy 1 AlalysleuLeuPProleuAspLysAspTyrTyrValValaArgLupProgly 17
 Db 52 GCTAAGCTGCTGCCGCTTGACAAAGACTACTAGCTGTCGCGACGACGCC 102
 RESULT 2
 LOCUS HSU08340 198 bp mRNA linear
 DEFINITION human clone NTK16 tyrosine kinase mRNA, partial cds.
 ACCESSION U08340
 VERSION U08340.1 GI:473879
 KEYWORDS human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1. (bases 1 to 198)
 AUTHORS Fuortes, M.
 TITLE Tumor Necrosis Factor and Adhesion: Joint Control of Neutrophil
 Activation
 JOURNAL Thesis (1994) Cell Biology and Anatomy, Cornell University,
 Graduate School of Medical Sciences
 REFERENCE 2 (bases 1 to 198)
 AUTHORS Fuortes, M.

TITLE Direct Submission
 JOURNAL Submitted (05-APR-1994) Michele Fuortes, Cell Biology and
 Anatomy/Medicine, Cornell University Medical College, 1300 York
 Avenue, New York, NY 10021, USA
 FEATURES source Location/Qualifiers
 1.198
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="NTK16"
 /cell_type="neutrophil"
 /tissue_type="blood"
 <1..>198
 /codon_start=1
 /evidence=not_experimental
 /product="tyrosine kinase"
 /protein_id="AA017743.1"
 /db_xref="GI:473880"
 /translation="DLARNILVESANHWKINDPGLAKLPLDKYVYVVERGCSPIF
 WYAPESLSDNIFRSQSDVWSFG"

BASE COUNT 41 a 66 c 52 g 39 t
 ORIGIN

Alignment Scores: 1.1e-08 Length: 198
 Pred. No.: 90.00 Matches: 17
 Score: 100.00% Conservative: 0
 Percent Similarity: 100.00% Mismatches: 0
 Best Local Similarity: 100.00% Indels: 0
 Query Match: 100.00% Gaps: 0
 DB: 9

US-09-397-967-15 (1-17) x HSU08340 (1-198)

Oy 1 AlalysleuLeuPProleuAspLysAspTyrTyrValValaArgLupProgly 17
 Db 67 GCTAAGCTGCTGCCGCTTGACAAAGACTACTAGCTGTCGCGACGACGCC 117
 RESULT 3
 LOCUS AX203080 448 bp DNA linear
 DEFINITION Sequence 1 from Patent W00152892.
 ACCESSION AX203080
 VERSION AX203080.1 GI:15392437
 KEYWORDS human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 448)
 AUTHORS Vasilos, G.
 TITLE Jak/stat pathway inhibitors and the uses thereof
 JOURNAL Patent: WO 0152892-A 1 26-JUL-2001;
 GENZYME CORPORATION (US)
 FEATURES source Location/Qualifiers
 1.448
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 BASE COUNT 79 a 150 c 127 g 92 t
 ORIGIN
 Alignment Scores: 2.63e-08 Length: 448
 Pred. No.: 90.00 Matches: 17
 Score: 100.00% Conservative: 0
 Percent Similarity: 100.00% Mismatches: 0
 Best Local Similarity: 100.00% Indels: 0
 Query Match: 100.00% Gaps: 0
 DB: 6
 US-09-397-967-15 (1-17) x AX203080 (1-448)
 Oy 1 AlalysleuLeuPProleuAspLysAspTyrTyrValValaArgLupProgly 17
 Db 306 GCTAAGCTGCTGCCGCTTGACAAAGACTACTAGCTGTCGCGACGACGCC 356

RESULT 4
AX203087 778 bp DNA linear PAT 30-AUG-2001
LOCUS AX203087
DEFINITION Sequence 8 from Patent WO0152892.
ACCESSION AX203087
VERSION AX203087.1 GI:15392445
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE
1 (bases 1 to 778)
Vasios, G.
JAK/stat pathway inhibitors and the uses thereof
Patent: WO 0152892-A 8 26-JUL-2001;
GENZYME CORPORATION (US)
FEATURES
source
1..778
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 131 a 265 c 219 g 163 t
ORIGIN
Alignment Scores:
Pred. No.: 4.75e-08 Length: 778
Score: 90.00 Matches: 17
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-09-397-967-15 (1-17) x AX203087 (1-778)
Oy 1 AlalysleuleuProleuasplysAsptyTtyValValargLupProgly 17
|||||
Db 187 GCTACGCTGCGCTTGACAAAGACTACTACTGCTGCGCGACGACGC 237
RESULT 5
AX203084 2914 bp DNA linear PAT 30-AUG-2001
LOCUS AX203084
DEFINITION Sequence 5 from Patent WO0152892.
ACCESSION AX203084
VERSION AX203084.1 GI:15392441
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE
1 (bases 1 to 2914)
Vasios, G.
JAK/stat pathway inhibitors and the uses thereof
Patent: WO 0152892-A 5 26-JUL-2001;
GENZYME CORPORATION (US)
FEATURES
source
1..2914
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 566 a 907 c 809 g 632 t
ORIGIN
Alignment Scores:
Pred. No.: 1.95e-07 Length: 2914
Score: 90.00 Matches: 17
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-09-397-967-15 (1-17) x AX203084 (1-2914)
Oy 1 AlalysleuleuProleuasplysAsptyTtyValValargLupProgly 17
|||||
Db 1856 GCTACGCTGCGCTTGACAAAGACTACTACTGCTGCGCGACGACGC 1906

RESULT 6
HSU31601 3620 bp mRNA linear PRI 24-OCT-1995
LOCUS HSU31601
DEFINITION Human tyrosine protein kinase (JAK3b) splice variant mRNA, complete cds.
ACCESSION U31601
VERSION U31601.1 GI:1039418
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE
1 (bases 1 to 3620)
Lai, K.S., Jin, Y., Graham, D.K., Witthuhn, B.A., Ihle, J.N. and
Lin, E.T.
A kinase-deficient splice variant of the human JAK3 is expressed in
hematopoietic and epithelial cancer cells
J. Biol. Chem. 270 (42), 25028-25036 (1995)
2 (sites)
Witthuhn, B.A., Silvennoinen, O., Mura, O., Lai, K.S., Cwik, C.,
Lin, E.T. and Ihle, J.N.
Involvement of the Jak-3 Janus kinase in signalling by interleukins
2 and 4 in lymphoid and myeloid cells
Nature 370 (6485), 153-157 (1994)
JOURNAL
MEDLINE
PUBMED
94294024
8022486
3 (bases 1 to 3620)
Liu, E.T.
Direct Submission
Submitted (14-JUL-1995) Edison T. Liu, Lindeberger Comprehensive
Cancer Center, University of North Carolina, Chapel Hill, NC
27599-7295, USA
Witthuhn, B.A. Nature 370, 153-157, 1994.
FEATURES
source
1..3620
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="19"
/map="19p12"
/clone="SKBR-3"
/cell_type="SKBR-3 breast cancer cells"
1..3620
/gene="JAK3b"
96..3380
/gene="JAK3b"
/function="tyrosine protein kinase"
/codon_start=1
/product="JAK3b"
/protein_id="AAC50226.1"
/db_xref="GI:1039419"
/translation="MAPSPEEPPLIPORSCSLSTPAGALHVLPPARGPQPTLSFS
FGDLAEDLCVOAKASGILPYHSILFALATEDSCMPESHFSVEDSTOVLVRI
RFTFPMNFGLEKRRGLRKDLASALDLPVLEHLPFAQRSDLSFVTRRRIRTRV
ECISLAVLDLARARQADQRPGLTKVSKACLPSPRLDIOGLSVTRRRIRTRV
RALRRVACQADNRSHLMARYIMDLERLDPAETGAEHVGLPGALSGHSDVGLRVAGD
GCIATMGQEQEVLQPCDEPEIVDISIKQAPRVGAGAEHLRVYTRDNOILAEKPG
LPEALSFVALVDGFRLLTDSQHFPEIVDISIKQAPRVGAGAEHLRVYTRDNOILAEKPG
GSRPQSYVLRSPQDPSFLITVQVONPGIDKKGLIRSPGTELLVSPHSLS
RELATLACMGDGLHVDVAVYITSCCIPRPREKSNLVQNGSPPTSLVQPOSQYL
SOMTEFKIPADSLMEHNLGHSFTKTYRCGRHEVVDGAEKTEVLAKWDAKHKCM
ESTLEASLSQVSRHLVYLHGVCNAGDSLVQEVVHLGADIVLKRGLHVPASMK
LOYVVKLALVNLLEDKGLPHGVNSARKVLLAREGADGSPPTKLDPGSPAVLSIE
MLTDRIIPWVAPFCLREAOITSLADKMGFGATVWEVSGVTPMISALDPAKIKQFED
ROOLPAKMTFELLALIQCCMAVEYVORSPRAYIRIDNSLISDSDELSPPTGALAP
RDLGMCAGQIYACODPTIFERRLAKYSQCGKNGFSGELCRDPTLGDNGALVAVKO
LOHSGPDQDQDFQRETOILKALHSDFTVTKRGVSYGGRGSLVMEYTPDSCGLRPL
QRRRARLDASRLLYSSQIKGMWYLGSRVCVRDLAARNLIVSEAHVA IADGLAK
LPLDXYVYVREPQSPITFWAPESLSDNIFRSQVMSFGVALLFLFYCDKSCSP
SAFLRMGCEBVDVPAICRLLELEEGORLIPADPACPAEVSAAGLASVSQSDVMAGVS
GKPAQA"

BASE COUNT 670 a 1152 c 1076 g 722 t
 ALIGNMENT SCORES:
 Pred. No.: 2,46e-07 Length: 3620
 Score: 90.00 Matches: 17
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 Gaps: 0

US-09-397-967-15 (1-17) x HSU31601 (1-3620)

OY 1 AlalysleuProleuAspLysAspTyrTyValValArgLupProgly 17
 Db 3006 GCTAAGCTGCTCCGCTTGACAAAGACTACTACGTGCTCCGAGCAGCC 3056

RESULT 7
 LOCUS AX203081 4064 bp DNA linear PAT 30-AUG-2001
 DEFINITION Sequence 2 from Patent WO0152892.
 ACCESSION AX203081
 VERSION AX203081.1 GI:15392438
 KEYWORDS
 SOURCE human
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 4064)
 AUTHORS Vassios, G.
 TITLE Jak/stat pathway inhibitors and the uses thereof
 JOURNAL GENZYME CORPORATION (US)
 FEATURES
 source Location/Qualifiers
 1..4064
 /organism="Homo sapiens"
 /db_xref="taxon:9606"

BASE COUNT 746 a 1292 c 1168 g 858 t

ALIGNMENT SCORES:
 Pred. No.: 2,78e-07 Length: 4064
 Score: 90.00 Matches: 17
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 Gaps: 0

US-09-397-967-15 (1-17) x AX203081 (1-4064)

OY 1 AlalysleuProleuAspLysAspTyrTyValValArgLupProgly 17
 Db 3006 GCTAAGCTGCTCCGCTTGACAAAGACTACTACGTGCTCCGAGCAGCC 3056

RESULT 8
 LOCUS HSU09607 4064 bp mRNA linear PRI 09-JUL-1994
 DEFINITION Human JAK family protein tyrosine kinase (JAK3) mRNA, complete cds.
 ACCESSION U09607
 VERSION U09607.1 GI:508730
 KEYWORDS
 SOURCE human
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 1124)
 AUTHORS Kawamura, M.; McGivicar, D.W.; Johnston, J.A.; Blake, T.B.; Chen, Y.;
 O'Shea, J.; Thompson, A.B.; Kelvin, D.; Stimples, J.E.; Ortaldo, D.R. and
 TITLE Molecular cloning of L-JAK, a Janus family protein-tyrosine kinase
 expressed in natural killer cells and activated leukocytes
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 91 (14), 6374-6378 (1994)
 MEDLINE 94291384

PUBMED 8022790
 REFERENCE 2 (bases 1 to 4064)
 O'Shea, J.
 TITLE Direct Submission
 JOURNAL Submitted (11-MAY-1994) John O'Shea, Leukocyte Cell Biology Section
 LEI BMD, National Cancer Institute FCRC, Bldg 560 Rm 3146 FCRC, Frederick, MD 21702, USA

FEATURES
 source Location/Qualifiers
 1..4064
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /cell_line="YF, HUT-78"
 /cell_type="natural killer cell, PHA activated T cells"
 /gene="JAK3"
 1..4064
 /gene="JAK3"
 96..3470
 /gene="JAK3"
 /codon_start=1
 /product="JAK family protein tyrosine kinase"
 /protein_id="AA19626.1"
 /db_xref="GI:508731"

CDS
 gene
 CDS

BASE COUNT 746 a 1292 c 1168 g 858 t

ALIGNMENT SCORES:
 Pred. No.: 2,78e-07 Length: 4064
 Score: 90.00 Matches: 17
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 Gaps: 0

US-09-397-967-15 (1-17) x HSU09607 (1-4064)

OY 1 AlalysleuProleuAspLysAspTyrTyValValArgLupProgly 17
 Db 3006 GCTAAGCTGCTCCGCTTGACAAAGACTACTACGTGCTCCGAGCAGCC 3056

RESULT 9
 LOCUS HSU70065 13562 bp DNA linear PRI 27-JAN-1997
 DEFINITION Human JAK3 gene, complete cds.
 ACCESSION U70065
 VERSION U70065.1 GI:1800224
 KEYWORDS
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 13562)
 AUTHORS Riedy, M.C.; Dutra, A.S.; Blake, T.B.; Modi, W.; Lal, B.K.; Davis, J.;
 Bosse, A.; O'Shea, J.J. and Johnston, J.A.
 TITLE Genomic sequence, organization, and chromosomal localization of

JOURNAL	human JAK3
GENOMICS	37 (1), 57-61 (1996)
PUBLISHED	97079659
REFERENCE	8921370
AUTHORS	2 (bases 1 to 13562)
TITLE	Riedy, M.C., Dutta, A.S., Blake, T.B., Modi, W., Lal, B.K., Davis, J., Bosse, A., O'Shea, J. J. and Johnston, J.A.
JOURNAL	Direct Submission
FEATURES	Submitted (09-SEP-1996) NTAMS, National Institutes of Health, 9000 Rockville Pike, Bethesda, MD 20892, USA
SOURCE	Location/Qualifiers
gene	1. .13562
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/chromosome="19"
	/map="19p13.1"
	201. .13562
	/gene="JAK3"
CDS	101in(201. .384,719. .842,1141. .1255,1462. .1607,2025. .2317,2865. .2987,3082. .3350,4264. .4450,6303. .6430,6630. .6761,7418. .7503,7918. .8044,8808. .8939,9019. .9171,9302. .9452,10839. .10978,11059. .11256,12009. .12125,12400. .12969)
	/gene="JAK3"
	/codon_start=1
	/product="JAK3"
	/protein_id="AAC50950.1"
	/db_xref="GI:1800225"
	/translation="MAPSEETPLIPQSCSLSTEAGALHVLIPARGGPORLSTFSRGYHMLDCEVQAAKAGILPVYHSLFALATEDSCFPPSHFISVEDASNOVLVYRISYFPPMFGLECHRGRLKRLDASALDIPVLEHLPAQHRSDISGRLPVGLSLKEOGEFLSLAVIDLARNARQARQARPGELLTYSYKAKLPRLDLOGLSVTRRAIRRYRRLPRAACADNRHSLMAKTYMDLERLDPAGAETHHVGLGALGSHDGLERFVAGDGIAMTQGEQVLAQPCDFPELVIDISLKOAPRAGPAEHRVYVTRDNOGLEFPGLPALSEVALAVDGFRLTTDSOHFECKEVAAPRLLEEVAEQCHGPIFLDEAFINKLKGSPSPGYSVILRRSPQDEPSFLITVCVOMLPEDPYGCLIRSGPTGFLVIGSRHSSLRELATCMGDGLHVDGVAVTTSCTTSCCIRPEKXSNLIYVQGHSPSTYVQPOSOYLSMTFPHKIPADSLFEMHENGHSFTKTYRGCGRHNVQGEARKTEVLLKYMDAKRNMESLLEASLSMSQSYRHLVTLHCGMACGDSITMOEPHGLCAIDMYLEKRGHLVYASNKLOVYKOLAVLNTLEBKGLPHGVNSAKRVLLARGADGSPFFIKLSDPGSAVLSLEMLTDRIIPWAPBECLEKQOTLSLEADKMGFGATWVEVSGVIMPLISALDPAPKALFYEDROOLPAKWTLELALLIOOCMAVEPVORPSRAVTRDNLNLSIDYEDLSDPTPALAPRDILMGAAQLYACODPTIEERHLKITYSQLGNGFSGVELCRDYPLDNTGALVAVQLOHSGDQOORDFOREIOTIKALHSDFLTVKRGVYSGGEPELRLVMEYLPBGGCLRDPLORRARRDASRLLLYSOICKMEYIGSRRCVHRDLAARNLIVSEAHVKIADSGLAKLPLDLDYVYVRRPGOSPTFWAPESLSDNIFSQSGVMSFGVLYLELFTYCDKSCSPSAFPLRMGCGRDVPALCRLELLEBQORLPAPPCAAEYHEMLKLCWADSPQDRPSPSALGPQDMLMSSSRGCEHTAFTAHPEGKHHSSFS"
exon	201. .384
	/gene="JAK3"
	/number=1
	385. .718
	/gene="JAK3"
	/number=1
	719. .842
	/gene="JAK3"
	/number=2
	843. .1140
	/gene="JAK3"
	/number=2
	1141. .1255
	/gene="JAK3"
	/number=3
	1256. .1461
	/gene="JAK3"
	/number=3
	1462. .1607
	/gene="JAK3"
	/number=4
	1608. .2024
	/gene="JAK3"
	/number=4
	2025. .2317
	/gene="JAK3"
intron	
exon	
intron	
exon	

intron	5	/number=5 2318..2864 /gene="JAK3"
exon	5	/number=5 2865..2987 /gene="JAK3"
intron	6	/number=6 2988..3081 /gene="JAK3"
exon	6	/number=6 3082..3350 /gene="JAK3"
intron	7	/number=7 3351..4263 /gene="JAK3"
exon	7	/number=7 4264..4450 /gene="JAK3"
intron	8	/number=8 4451..6302 /gene="JAK3"
exon	8	/number=8 6303..6630 /gene="JAK3"
intron	9	/number=9 6631..6629 /gene="JAK3"
exon	9	/number=9 6630..6761 /gene="JAK3"
intron	10	/number=10 6762..7417 /gene="JAK3"
exon	10	/number=10 7418..7503 /gene="JAK3"
intron	11	/number=11 7504..7917 /gene="JAK3"
exon	11	/number=11 7918..8044 /gene="JAK3"
intron	12	/number=12 8045..8807 /gene="JAK3"
exon	12	/number=12 8808..8939 /gene="JAK3"
intron	13	/number=13 8940..9018 /gene="JAK3"
exon	13	/number=13 9019..9171 /gene="JAK3"
intron	14	/number=14 9172..9301 /gene="JAK3"
exon	14	/number=14 9302..9452 /gene="JAK3"
intron	15	/number=15 9453..10839 /gene="JAK3"
exon	15	/number=15 10839..10978 /gene="JAK3"
intron	16	/number=16 10979..11058 /gene="JAK3"
exon	16	/number=16 11059..11256 /gene="JAK3"
exon	17	/number=17

```

        intron          11257..12008
                        /gene="JAK3"
                        /number=17
        exon            12009..12125
                        /gene="JAK3"
                        /number=18
        intron          12126..12399
                        /gene="JAK3"
                        /number=18
        exon            12400..13562
                        /gene="JAK3"
                        /number=19

BASE COUNT      2939 a      3476 c      4009 g      3131 t      7 others
ORIGIN
Alignment Scores:
Pred. No.:      1.01e-06      Length:      13562
Score:          90.00      Matches:      17
Percent Similarity: 100.00%      Conservative: 0
Best local Similarity: 100.00%      Mismatches: 0
Query Match:      100.00%      Indels:      0
DB:              9      Gaps:      0

US-09-397-967-15 (1-17) x HSU70065 (1-13562)
OY      1      AlAlaYsLeuLeuProLeuAspLysAspTyrValValAlaArgLupProGly 17
        |||||
Db      12505 GCTAAGCTGCTGCCGCTTGACAAAGACTACTACGTGCTGCCGAGCCAGGC 12555

RESULT 10
AF513860      21323 bp      DNA      linear      PRI 09-JUL-2002
LOCUS      Homo sapiens Janus kinase 3 (a protein tyrosine kinase, leukocyte)
DEFINITION      (JAK3) gene, complete cds.
ACCESSION      AF513860
VERSION      AF513860.1      GI:21263105
KEYWORDS
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 21323)
AUTHORS      Rieder,M.J., Armel,F.Z., Carrington,D.P., Ozuna,M., Kuldanek,S.A.,
        Rajkumar,N., Toch,E.J., Yi,Q. and Nickerson,D.A.
        Direct Submission
        Submitted (20-MAY-2002) Genome Sciences, University of Washington,
        1705 NE Pacific, Seattle, WA 98195, USA
        To cite this work please use: SeattleSNPs, NHLBI HUG6682 Program
        for Genomic Applications, UW-PHRC, Seattle, WA (URL:
        http://pga.gs.washington.edu).
FEATURES
        source
            1..21323
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
        repeat_region
            1..214
                /rpt_family="Alu"
        variation
            290
                /rpt_type=dispersed
                /frequency="0.13"
                /replace="a"
        variation
            370
                /frequency="0.02"
                /replace="t"
        variation
            575
                /frequency="0.33"
                /replace="a"
        variation
            622
                /frequency="0.01"
                /replace="a"
        repeat_region
            673..965
                /rpt_family="Alu"
                /rpt_type=dispersed
        variation
            681

        /frequency="0.02"
        /replace="a"
        784
        /frequency="0.13"
        /replace="a"
        925
        /frequency="0.33"
        /replace="g"
        970
        /frequency="0.01"
        /replace="t"
        994
        /frequency="0.96"
        /replace="n"
        1019..1270
        /rpt_family="Alu"
        /rpt_type=dispersed
        1039
        /frequency="0.04"
        /replace="c"
        1051
        /frequency="0.01"
        /replace="a"
        1061
        /frequency="0.06"
        /replace="t"
        1064
        /frequency="0.04"
        /replace="t"
        1273..1336
        /rpt_family="Alu"
        /rpt_type=dispersed
        1337..1630
        /rpt_family="Alu"
        /rpt_type=dispersed
        1467
        /frequency="0.30"
        /replace="a"
        1513
        /frequency="0.04"
        /replace="t"
        1564
        /frequency="0.23"
        /replace="a"
        1590..2091
        /note="Region not scanned for variation"
        1631..1706
        /rpt_family="Alu"
        /rpt_type=dispersed
        1966..2274
        /rpt_family="Alu"
        /rpt_type=dispersed
        gene
            <2359..20650
                /gene="JAK3"
                /join(<2359..2557,2891..3014,3300..3411,3619..3764,
                join(2374..2557,5029..5151,5245..5402,6455..6566,7133..7319,
                4181..4475,5029..5151,5245..5402,6455..6566,7133..7319,
                8406..8533,8733..8864,9583..9667,10746..10873,
                11582..11714,11794..11945,12076..12226,13868..14007,
                14089..14278,14999..15123,15397..15569,16178..16295,
                16580..16690,19892..20650)
                /gene="JAK3"
        /product="Janus kinase 3 (a protein tyrosine kinase,
        leukocyte)"
        join(2374..2557,2891..3014,3300..3411,3619..3764,
        4181..4475,5029..5151,5245..5402,6455..6566,7133..7319,
        8406..8533,8733..8864,9583..9667,10746..10873,
        11582..11714,11794..11945,12076..12226,13868..14007,
        14089..14278,14999..15123,15397..15569,16178..16295,
        16580..16690,19892..20059)
        /gene="JAK3"
        /codon_start=1
        /product="Janus kinase 3 (a protein tyrosine kinase,
        leukocyte)"

```



```

/protein_id="AA044860.1"
/db_xref="GI:21263106"
/translation="MAPSEETPLIPORSCSLSTEWAGLHVLPARGGPPORLSFS
RGDHLAEDLCNOAKASGILPYHSLFLATPDLSCMPRSSFSVEASTOYLXRI
RPFPMFEGLEKCHRGKRLKDLASILDLPVLEHLPQHRSDLVSGRLPVGSLKEQG
ECLSLAVLDLARMAREQARPEGLKTLVSKCLPSLRLDILQSLFVTRRIIRTVR
RALRVACQADRHSLIMKYIMDLERLPAGAEFHVGLPGALGHDGLLRVAG
GCIANTQGEVLOPFCDFPEIVDISIKQAPVAGEHRLVTRTNOILAEFPG
LPEALSYVALVDGYPRLTDSQHFCKEVAAPRLLEVEQCHGPTLDEAIKRLKGTG
GSRGCSYLRSPDRDPDFSLTVCVONPGRPYKGLIRSPGTGRLVYGLSRPSL
RELLATCMDCGLHVDGAVTLTSCCIPRKEKSNLIVORGSPTSSLVQSQOYL
SMTFHKIPADSLHNVHNGHSTFKITRGCHREVDGFAKTEVLKVMKHNKM
ESFLASLMSQVSYRHVLHGVCMAGDSTWQEFVHLGALIDMLRKRGHLVPSMK
LOVYQOLAYALNYLEDKGLPHGNVSAKRYLAREGADGSPFIKLSDEGSPAVLSLE
MLTDRIIPWAPCELEAQTLSLEADKMGGAIVWEVSGVTMPSALDPARKLOFED
ROOLPARKMTETALLIOCMAYEPVORSEFRAVITDINLSISDVELLSDTPRGALAP
RDGLMNGALYVACOPTIFEEHRLKYSIQSGKNGSVYELCRDPLGDTWGLVAVKQ
LOHSGPDDORDFOREIQTILKALHSDPIYAKVRYSGPGQSLRLWEYVLPSCGLRDL
OHNHARLDAISRLYSQICKGMEYLGSRVCHRODLAANNILVESEAHVKIADFGIAK
LPLDKDYVVRPEGQSPIFWYAPESLSNIPRSROSDVMSFGVLYELFTYCDKSCSP
SAEFLRMGCEKRDVPALCRFLLEEGORLPAPPCAPVHMLKLCVAPSPQDRPSF
SALGPQLMLMGSGRCETHAFTAHPECKHNSLFS"
2579

```

variation

variation

variation

repeat_region

variation

variation

variation

variation

variation

repeat_region

variation

variation

variation

variation

```

/gene="JAK3"
/frequency="0.06"
/replace="a"
3385
/gene="JAK3"
/frequency="0.06"
/replace="a"
3442
/gene="JAK3"
/frequency="0.04"
/replace="t"
3451

```

variation

Alignment Scores:

```

Pred. No.: 1 64e-06 Length: 21323
Score: 90.00 Matches: 17
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

```

US-09-397-967-15 (1-17) x AF513860 (1-21323)

```

QY 1 AlalysleuLeuProleuAsplysAspTyTrValValArgGluProGly 17
Db 15502 GCTAACCTGCTGCCGCTTGACAAAGACTACTACGTGCTGCCGACGACGCG 15552

```

RESULT 11

AC007201/c

LOCUS

DEFINITION

AC007201

AC007201.1

KEYWORDS

HTG.

SOURCE

ORGANISM

REFERENCE

AUTHORS

lamerdin,J.E., McCready,P.M., Skowronski,E., Scott,D.,
Stilwagen,S., Phan,H., Velasco,N., Do,L., Regala,W., Terry,A.,
Dangnan,L., Erler,A., Christensen,M., Georgescu,A., Avila,J.,
Atlix,C., Andreise,T., Amico-Keller,G., Coefield,J., Duarte,S.,
Lucas,S., Bruce,R., Thomas,P., Quan,G., Krommiller,B., Arellano,A.,
Sanders,C., Ow,D., Nolan,M., Trong,S., Kobayashi,A., Olsen,A.S. and
Carrano,A.V.

Sequence analysis of a 5.7 Mb region in 19p13.1

Unpublished

2 (bases 1 to 41006)

REFERENCE

AUTHORS

JOURNAL

COMMENT

Submitted (03-APR-1999) Joint Genome Institute, Lawrence Livermore
National Laboratory, 7000 East Ave., Livermore, CA 94551, USA
Map and sequence oriented from p telomere to centromere. Cosmid
R34383 overlaps cosmid F19847 (AC005952) to the left from bases 1
to 2,923 of this accession, and overlaps cosmid R31408 (AC005796)
to the right from bases 37,823 to 41,006. Additional chromosome 19
map and sequence information are available at:
<http://www-bio.livnl.gov/db/rp/genome/genome.html>.

FEATURES

source

1. 41006

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="19"

/map="19p13.1 between JAK3 and PDE4C"

/clone="R34383"

/cell_line="5HL2-B"

/clone.lib="L19NC03 R chromosome 19-specific cosmid library"

/note="Cosmid library constructed at LNL from flow-sorted
chromosomes from human-hamster hybrid 5HL2-B, which
carries chromosome 19 as its only human chromosome."

```

repeat_region      1. .67
/rpt_family="AluY/monomer"
repeat_region      143. .441
/rpt_family="AluYb"
repeat_region      complement(578. .885)
/rpt_family="AluXs"
CDS                complement(<1078. .1272)
/note="LEYDIG INSULIN-LIKE PEPTIDE PRECURSOR"
/codon_start=1
/product="INL3 HUMAN [AA 1- 65]"
/db_xref="GI:4558770"
/translation="MDPRLPAWALVLLGPALVAFALGAPAPPEMRKELCGHHFVALVR
VCGGPRMSTEARAPATGCDRE"
repeat_region      complement(1486. .1786)
/rpt_family="AluXs"
repeat_region      complement(1803. .1970)
/rpt_family="AluSg/X"
repeat_region      complement(1975. .2275)
/rpt_family="AluYb"
repeat_region      complement(2279. .2444)
/rpt_family="AluYo"
repeat_region      complement(2447. .2747)
/rpt_family="AluXs"
repeat_region      complement(2748. .2877)
/rpt_family="AluYo"
repeat_region      3052. .3357
/rpt_family="AluXs"
repeat_region      complement(3365. .3412)
/rpt_family="MIR"
repeat_region      complement(3773. .3846)
/rpt_family="(GA)n"
repeat_region      complement(3851. .4142)
/rpt_family="AluXs"
repeat_region      complement(4914. .4995)
/rpt_family="MIR"
repeat_region      5386. .5686
/rpt_family="AluY"
repeat_region      5739. .5848
/rpt_family="FLAM_A"
repeat_region      5870. .6170
/rpt_family="AluXs"
gene               complement(5918. .24239)
/note="JAK3"
/note="Janus kinase 3 (a protein tyrosine kinase,
leukocyte)"
mRNA               complement(join(5918. .6676,9874. .9984,10269. .10386,
10994. .11166,11440. .11564,12285. .12474,12556. .12695,
14337. .14487,14618. .14769,14849. .14981,15690. .15817,
16885. .16979,17698. .17829,18029. .18156,19243. .19429,
19986. .20107,21155. .21312,21406. .21528,22082. .22376,
22793. .22938,23146. .23257,23543. .23666,24000. .24239))
/gene="JAK3"
/product="Human JAK family protein tyrosine kinase (JAK3)
mRNA, complete cds."
repeat_region      complement(6416. .6472)
/rpt_family="MIR"
CDS                complement(join(6509. .6676,9874. .9984,10269. .10386,
10994. .11166,11440. .11564,12285. .12474,12556. .12695,
14337. .14487,14618. .14769,14849. .14981,15690. .15817,
16885. .16979,17698. .17829,18029. .18156,19243. .19429,
19986. .20107,21155. .21312,21406. .21528,22082. .22376,
22793. .22938,23146. .23257,23543. .23666,24000. .24239))
/gene="JAK3"
/note="TYROSINE-PROTEIN KINASE JAK3; JANUS KINASE 3;
JAK-3; LEUKOCYTE JANUS KINASE; L-JAK"
/codon_start=1
/product="JAK3_HUMAN"
/protein_id="A02274.1"
/db_xref="GI:4558771"
/translation="MAPSEETPLIPORSCSLSTEAGALHVLIPARGSPORLSPS
FGHLADLCVQAKASGILPYVHSIALATEDISCFPPSHITSVEDASQVLLIRI
RFYFPMWFGLEKCHRFGLRKDLASALLDPLVLEHLFAOHNSDLVSGRLPVGLSLKEOG

```

```

repeat_region      complement(6747. .7040)
/rpt_family="AluXs"
repeat_region      complement(7111. .7412)
/rpt_family="AluSg"
repeat_region      7441. .7688
/rpt_family="AluSp"
repeat_region      complement(7985. .8020)
/rpt_family="(CAAAn)"
repeat_region      complement(8021. .8315)
/rpt_family="AluYb"
repeat_region      complement(8316. .8454)
/rpt_family="AluYb"
repeat_region      8499. .8564
/rpt_family="LINE2"
repeat_region      8643. .8902
/rpt_family="AluYo"
repeat_region      complement(8929. .9224)
/rpt_family="AluXs"
repeat_region      9243. .9536
/rpt_family="AluYo"
repeat_region      complement(9560. .9753)
/rpt_family="MIR"
repeat_region      11759. .12044
/rpt_family="AluXs"
repeat_region      complement(12074. .12154)
/rpt_family="MIR"
repeat_region      13194. .13323
/rpt_family="FLAM_A"
repeat_region      13337. .13638
/rpt_family="AluXs"
repeat_region      complement(13651. .13805)
/rpt_family="AluYb"
repeat_region      13832. .14140
/rpt_family="AluXs"
repeat_region      complement(16047. .16337)
/rpt_family="AluXs"
repeat_region      complement(17180. .17480)
/rpt_family="AluXs"
repeat_region      complement(17487. .17609)
/rpt_family="FLAM_C"
repeat_region      complement(17910. .17976)
/rpt_family="MIR"
repeat_region      complement(18301. .18479)
/rpt_family="AluSg/X"
repeat_region      complement(18480. .18780)
/rpt_family="AluXs"
repeat_region      complement(19505. .19804)
/rpt_family="AluSc"
repeat_region      20248. .20548
/rpt_family="AluY"
repeat_region      complement(20590. .20882)
/rpt_family="AluYo"
repeat_region      23362. .23462
/rpt_family="MIR"
repeat_region      complement(24063. .24094)
/rpt_family="GC-rich"

```

repeat_region 24283..24582
 /rpl_family="Alusx"
 repeat_region 24583..24608
 /rpl_family="(GAA)n"
 repeat_region 24851..24930
 /rpl_family="Alu/FLAM"
 repeat_region complement(24932..25220)
 /rpl_family="Alusx"
 repeat_region 25221..25284
 /rpl_family="Alu/FLAM"

Alignment Scores:

Pred. No.: 3.29e-06 Length: 41006
 Score: 90.00 Matches: 17
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0

US-09-397-967-15 (1-17) x AC007201 (1-41006)

Qy 1 AlalysleuProleuAspLysAspTyrTyValValArgGluProgly 17

Db 11061 GCTAAGCTGCTGCCGCTTGACAAAGACTACTACGTGCTCCGACGACGAGC 11011

RESULT 12

MUSJAK3A

LOCUS MUSJAK3A 3528 bp mRNA linear ROD 24-APR-1996
 DEFINITION Mus musculus JAK3 gene, complete cds.
 ACCESSION L40172.1 GI:1019909
 VERSION L40172.1 GI:1019909
 KEYWORDS JAK3 gene.

SOURCE Mus musculus (strain BALB/c x 129 F2) (clone: clone-s12-18) neonate
 thymus CDNA to mRNA.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 3528)
 Gurliak, C.B. and Berg, L.J.

AUTHORS Murine JAK3 is preferentially expressed in hematopoietic tissues
 and lymphocyte precursor cells
 JOURNAL Blood 87 (8), 3151-3160 (1996)
 MEDLINE 96184772
 PUBMED 8605329

FEATURES

SOURCE

Location/Qualifiers

1..3528
 /organism="Mus musculus"
 /strain="BALB/c x 129 F2"
 /db_xref="taxon:10090"
 /clone="clone-s12-18"
 /cell_type="T cell precursor"
 /tissue_type="thymus"
 /dev_stage="neonate"
 1..3528
 /gene="JAK3"
 1..71
 /gene="JAK3"
 72..3374
 /gene="JAK3"
 /codon_start=1
 /protein_id="AAC2085.1"
 /db_xref="GI:1019910"

CDS

/translation="MAPPSEETPLIPORSCSLSSSEAGALHVLPPRGCPORLIFS
 EGDYLAEDLCVRAKACGILPVYHSLFALATEDFCWFPSPSHFIEDVDQVLYRL
 RYTPDWFGLTCHRGKRLDLSALDLHLVLEFAOHSRLVSGRLPVGSLMKRGG
 EFLSLAVLDLQAMAREQACRQGLKLTYSYKALPLPSLRVIOGQNFVRRIRRYV
 LALRRVACQADRALMAKYILDLERLHPAATETETRGVPGQEPGGLRVAAGDGI
 SMSGQDELFOFCADPEIVDVISIKQAPRGAGERLVTVMDSHLLAEFGPIPE
 PGTYILRRSPQDDSTFLTRACVTPPLGPKGCLIQDPSGAFSLVGLSQPHRSRL
 LAACWMSGLVDAALNLVSCAPRPEKSNLIVRRGCTPAPPCGSSCCALQOLS
 FHTIPDLSFMENLGHGFTKIFGRREEVVCGETHDSEVLKLVDSRRKCMSEFL
 EASLMSOVSYPHLVLLHGVCMAGDSIMVQEFVYLGALIDMYLRKHGHVYSAMKLOVT

3'UTR
 polyA-signal
 BASE COUNT 648 a 1104 c 1036 g 740 t
 ORIGIN

Alignment Scores:

Pred. No.: 5.63e-06 Length: 3528
 Score: 83.00 Matches: 16
 Percent Similarity: 94.12% Conservative: 0
 Best Local Similarity: 94.12% Mismatches: 1
 Query Match: 92.22% Indels: 0
 DB: 10 Gaps: 0

US-09-397-967-15 (1-17) x MUSJAK3A (1-3528)

Qy 1 AlalysleuProleuAspLysAspTyrTyValValArgGluProgly 17

Db 2970 GCTAAGCTGCTGCCCTCGAAGAGACTACTACGTGCTCCGACGACGCTGGC 3020

RESULT 13

MUSPYKIN

LOCUS MUSPYKIN 3723 bp mRNA linear ROD 21-SEP-1994
 DEFINITION Mouse protein tyrosine kinase (JAK3) mRNA, complete cds.
 ACCESSION L32955
 VERSION L32955.1 GI:529238

KEYWORDS Interleukin 2; Interleukin 4; Janus kinase; protein tyrosine
 kinase.
 SOURCE Mus musculus (strain BALB/c, sub-species domesticus) CDNA to mRNA.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 3723)
 Witthuhn, B.A., Silvennoinen, O., Mura, O., Lai, K.S., Witk, C.,
 Liu, E.T. and Ihle, J.N.
 TITLE Involvement of the Jak-3 Janus kinase in signalling by Interleukins
 2 and 4 in lymphoid and myeloid cells
 JOURNAL Nature 370 (6485), 153-157 (1994)
 MEDLINE 94294024
 PUBMED 8022486

COMMENT This entry has been reported under the accession number L329555 in
 the Nature article.
 FEATURES
 source
 Location/Qualifiers

1..3723
 /organism="Mus musculus"
 /strain="BALB/c"
 /sub_species="domesticus"
 /db_xref="taxon:10090"
 /cell_line="Abelson transformed pre-B-cell line 22D6"
 1..3723
 /cell_type="lymphoid and myeloid"

CDS

1..12
 /gene="jak3"
 13..3312
 /gene="jak3"
 /standard_name="Janus kinase, mjak3"
 /codon_start=1
 /evidence="experimental"
 /product="protein tyrosine kinase"
 /protein_id="AA21565.1"
 /db_xref="GI:529239"
 /translation="MAPPSEETPLIPORSCSLSSSEAGALHVLPPRGCPORLIFS
 KOLAYALNTEDEKGLPHGNVSAKRVLLAREGQDNPPIKISDPGVSPFYLSLEMLTD
 RIPWVAPEDIOEORTLCLEADKMGKATTEVFSSGPAHITSLSPAKLKRYEPQGL
 PALKMTLGLITQCMAYDPGRPSFRAILDNLGLTISDELSDIPGIPSPRDEL
 CGCAQIYACQDPAITEERHLKTYISLGLNGFSGVELCYRDLGNTGPIVAVKQLOHS
 GPDQDFOREIOIILKALHSDFIVKRGVSGVSCSLRLVMEVLPSCGLDFLOHR
 ARLHTDRLLEFAMQICKMEYLGARCVHRLAARNTLVESEAIKTIADPCLATPL
 GKDYVVERPGSPIFWYAPESLSNIFRSQSDVMSRGVLYELFTYCDKSCSFAEP
 LRMGPEREGPCLRLLELAGRLLPRPPICPIEVDELMLQCMAPSHDRPARFTLS
 POLDALMRGPG"

```

RDGYLAEDICVBARAKACALIPYVHSILFALATEDFSCMPSPSHIFCIEDVDIOVLVYRL
REFEPMFLETCRFRGLKRDLSALIDHLVLEHLEFAOHRSDVSGRLPVGSGMEKOG
EFLSLAVLADLAQMARQAOAPRPGELTKTVSKACLPSEADYIOGONFVRRIRRTVYV
LALLPGCRLPGRPVYALMAY11DLERLHPAATTEFRVGLPEAOEPGLLRACNGI
PMSNNELFQTECDPEIIVDSINQAPRVGPEHRLVVTMDGHLLEAFGLPEA
LSVVALVDGYFLICDSRHYPCKEVAPRPLEEADVCHGPTTLDFAIHKLKAAASLP
GVYILRRPODYSPLITACVOTPLGPDYKGLIROPDGSFSLVGLSQPHRSRELL
AACNSGIRADGALYITSCAPRKEKSNLIYVRGDNAPAPGSCSPSCALITLSE
HTTPSISGIRHNLGHGSEFTKIPRGSRRVDEGTHDSVLLKMDSRHNCMESFLE
AASLMGOSTPHILVLLHAGCMAGDSIMOEFTVYLAIDMYLRRGHILVASMKLOVTK
OLAYALNLEDKCLPHGNVSARKVYLARBGGRGPAHITSLERPAKXIKFYEDQOLP
IPWAPRPELOEAOITCLPQWGFATWVFORGPAHITSLERPAKXIKFYEDQOLP
ALMWTLAGLITOCMAVDGRPSFRAILDLNGLITSDTELDPTPOTIPSPRBDIC
VAGQIYACODPAIFEBERHLKXISLKGNGFSGVLCRDPLDNTGPIVAVQLOHS
VPOQDRFOREIOLKALHSDPIYKRVGYSGRQSLRLVMEYLPSGCLRLDLLOHR
GLHTRDRLIFAMQICKMEYLGARCVHRLARNLIVSEAHVKIADGSLAKLPLG
KDYVVERPGOSPFMWYAPRSLSDNIFRSQSDVMSFGVYELFTYCDSCSPSAEFL
RMKGPERGPPICRLLELLAEGRLRPPPTCTEVEQLMQLCVAPEPHDRPAPATLSP
QLDPLMRGRPG"
/mat_peptide
13..3309
/gene="jak3"
/product="protein tyrosine kinase"
/standard_name="Janus kinase, mujak3"
/evidence=experimental
3'UTR
3313..3723
/gene="jak3"
polyA_site
3723
/gene="jak3"
BASE COUNT 711 a 1162 c 1079 g 771 t
ORIGIN

Alignment Scores:
Pred. No.: 5.96e-06 Length: 3723
Score: 83.00 Matches: 16
Percent Similarity: 94.12% Conservative: 0
Best Local Similarity: 94.12% Mismatches: 1
Query Match: 92.22% Indels: 0
Gaps: 0
DB: 10

US-09-397-967-15 (1-17) x MUSPTYKIN (1-3723)
OY 1 AlalysleuLeuProleuAspTyTyTyValaValArgGluProGly 17
|||||
Db. 2908 GCTAACTGCTGCTCCCTGGGAAGACTACTAGTGTCCGCGAGCCTGGC 2958

RESULT 14
MUSJAK3H 4016 bp mRNA linear ROD 13-SEP-1994
LOCUS Mus musculus (clone 32D5-1) protein tyrosine kinase (JAK3) mRNA.
DEFINITION L33768
ACCESSION L33768.1 GI:508533
VERSION protein tyrosine kinase.
KEYWORDS Mus musculus cDNA to mRNA.
SOURCE Mus musculus
ORGANISM Mus musculus; Chordata; Craniata; Vertebrata; Euteleostomi;
Eukaryota; Metazoa; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 4016)
AUTHORS Rane,S.G. and Reddy,E.P.
TITLE JAK3: a novel JAK kinase associated with terminal differentiation
of hematopoietic cells
JOURNAL Oncogene 9 (8), 2415-2423 (1994)
MEDLINE 94309920
PUBMED 7518579
FEATURES
source 1..4016
location/Qualifiers
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="32D5-1"
/cell_line="32Dc13"
/tissue_lib="lambda gtl1"
7..3906
/note="protein tyrosine kinase"
/codon_start=1

mat_peptide
13..3309
/gene="jak3"
/product="protein tyrosine kinase"
/standard_name="Janus kinase, mujak3"
/evidence=experimental
3'UTR
3313..3723
/gene="jak3"
polyA_site
3723
/gene="jak3"
BASE COUNT 711 a 1162 c 1079 g 771 t
ORIGIN

Alignment Scores:
Pred. No.: 5.96e-06 Length: 3723
Score: 83.00 Matches: 16
Percent Similarity: 94.12% Conservative: 0
Best Local Similarity: 94.12% Mismatches: 1
Query Match: 92.22% Indels: 0
Gaps: 0
DB: 10

US-09-397-967-15 (1-17) x MUSPTYKIN (1-3723)
OY 1 AlalysleuLeuProleuAspTyTyTyValaValArgGluProGly 17
|||||
Db. 2908 GCTAACTGCTGCTCCCTGGGAAGACTACTAGTGTCCGCGAGCCTGGC 2958

RESULT 14
MUSJAK3H 4016 bp mRNA linear ROD 13-SEP-1994
LOCUS Mus musculus (clone 32D5-1) protein tyrosine kinase (JAK3) mRNA.
DEFINITION L33768
ACCESSION L33768.1 GI:508533
VERSION protein tyrosine kinase.
KEYWORDS Mus musculus cDNA to mRNA.
SOURCE Mus musculus
ORGANISM Mus musculus; Chordata; Craniata; Vertebrata; Euteleostomi;
Eukaryota; Metazoa; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 4016)
AUTHORS Rane,S.G. and Reddy,E.P.
TITLE JAK3: a novel JAK kinase associated with terminal differentiation
of hematopoietic cells
JOURNAL Oncogene 9 (8), 2415-2423 (1994)
MEDLINE 94309920
PUBMED 7518579
FEATURES
source 1..4016
location/Qualifiers
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="32D5-1"
/cell_line="32Dc13"
/tissue_lib="lambda gtl1"
7..3906
/note="protein tyrosine kinase"
/codon_start=1

misc_feature
2199..3024
/note="JAK homology domain II (JH2)"
3076..3876
/note="JAK homology domain I (JH1)"
BASE COUNT 742 a 1203 c 1235 g 836 t
ORIGIN

Alignment Scores:
Pred. No.: 6.47e-06 Length: 4016
Score: 83.00 Matches: 16
Percent Similarity: 94.12% Conservative: 0
Best Local Similarity: 94.12% Mismatches: 1
Query Match: 92.22% Indels: 0
Gaps: 0
DB: 10

US-09-397-967-15 (1-17) x MUSJAK3H (1-4016)
OY 1 AlalysleuLeuProleuAspTyTyTyValaValArgGluProGly 17
|||||
Db. 3544 GCTAACTGCTGCTCCCTGGGAAGACTACTAGTGTCCGCGAGCCTGGC 3594

RESULT 15
MMU71201 10326 bp DNA linear ROD 15-JAN-1997
LOCUS Mus musculus protein tyrosine kinase JAK3 gene, complete cds.
DEFINITION U71201
ACCESSION U71201
VERSION U71201.1 GI:1178748
KEYWORDS Mus musculus.
SOURCE Mus musculus.
ORGANISM Mus musculus; Chordata; Craniata; Vertebrata; Euteleostomi;
Eukaryota; Metazoa; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 10326)
AUTHORS Kumar,A., Toscani,A., Rane,S. and Reddy,E.P.
TITLE Structural organization and chromosomal mapping of JAK3 locus
JOURNAL Oncogene 13 (9), 2009-2014 (1996)
MEDLINE 8934548
PUBMED 8934548
REFERENCE 2 (bases 1 to 10326)
AUTHORS Kumar,A. and Reddy,E.P.
TITLE Direct Submission
JOURNAL Submitted (18-SEP-1996) Fels Institute, Temple University, 3307 N.
Broad St, Rm 339 AHB, Philadelphia, PA 19140, USA
FEATURES
source 1..10326
location/Qualifiers
/organism="Mus musculus"
/strain="129/J"

```

mrna
/db_xref="taxon:10090"
/chromosome="19"
/map="19p13.1-p13.2"

3010(92) 287,441..564,740..851,963..1108,1317..1599,
2170..2289,2372..2529,2617..2728,3230..3413,3781..3902,
3998..4129,4596..4680,5022..5149,5642..5774,5905..6056,
6177..6327,6856..6995,7074..7260,7427..7551,7629..7801,
8524..8641,8907..9017,9707..10212)
/product="JAK3"

3010(104) 287,441..564,740..851,963..1108,1317..1599,
2170..2289,2372..2529,2617..2728,3230..3413,3781..3902,
3998..4129,4596..4680,5022..5149,5642..5774,5905..6056,
6177..6327,6856..6995,7074..7260,7427..7551,7629..7801,
8524..8641,8907..9017,9707..9772)
/function="protein tyrosine kinase"
/codon_start=1
/product="JAK3"
/protein_id="AAB40917.1"
/db_xref="GI:1778749"

/translation="MAPSEETPLIPQSCSLSSSEAGALHVLPPRGPPQRLSFS
FGDYLAEDLCVRAKACGILPVYHSLFALATEDSCWEPSSHIFCIEDVDTQVLYRL
RFYEPDMFGLETCRFGRLKDLTSAILDLHLFPAOHRSDLVSGRLPYGLSMKEDG
EFLSLAVLDLAOMAREQARQREGELTKVSYKACLPISBDVIGQNFVTRRLIRTV
LALRWSPARPTAARAGOVYLDLERLHPATTTTFRYGLDPAQEEPGLLRVAGDNGIS
MSSGDDELFTQFCDPRFIVSVSTKOPTGSGREHRLVTVTRMDGHTLEAFPGLEAL
SPVALNDGVRLLCDSDRHRYCKEVAAPRLLEEAEALHGHTLDFAIHRLKAASLPG
TYLRNSPDYDSFLTACVOTPLGPDYKGLITRODPSGAFSLVGLSPTEAGCTLO
CWNSSGLVDDAALNLTSCCAPRPEKSNLIVRRCTPAPAPGCSPPSCALNOLSPHT
IPDLSLGHENLGHGSFTKIFGRRRREVVDGETHDESVLLKVMDSRRHNCMESLEAS
LMSQVSPHLVLHGVCMAGDSIMVOEYVIGAIDMYLRKHIVASAKLQVTKOLA
YALNYLEDKGLPHGNVSARKVILLAREGGDGNPFITLSDPGVSPVLSLEMLDRIIPW
VAPECLQEAQTLCLLEADKNGFGATTWVEVSGCPAHITSLLEPAKKLFYEDQQLPALK
WTELAGLITQCMAYDPGRNRSFRATLRDLNGLITSDYELLSDPTGIPSPRDELGPAK
QIYACODPAIFEEHRLKYISLKGKNGFSELCRYDPLDNTGPLVAVKQLQHSQPDQO
RDRREIQIILKALHSDFIKRVGVSYPGROSRLIWMEXLPSGCLRDFLQHRALHVT
DRLLFAWQICKMEYLGAARCVCYHDLARNITLVESEAHVKIADFGIAKLILPLGKDY
VVEREGOSPFIWAPESLSDNIFRSQSDVMSPGVLYELFTYCDKSCSPSAEFLRMNG
PEREGPPLRLELLELAEGRRLPPPTCPTEVOELMOLICVAPSRITGQPSAP"

BASE COUNT 2088 a 3007 c 2880 g 2300 t 51 others

Alignment Scores:

Prod. No.: 1.78e-05 Length: 10326
Score: 83.00 Matches: 16
Percent Similarity: 94.12% Conservative: 0
Best Local Similarity: 94.12% Mismatches: 1
Query Match: 92.22% Indels: 0
DB: 10 Gaps: 0

US-09-397-967-15 (1-17) x MMU71201 (1-10326)

OY 1 AlAlYsLeuDeuPProLeuAspTYTYVaiIaIrgLupProGly 17
DB 7734 GCTAAGCTGCTGCCCTGGGAAAGACTACTAGCTGCTCCGAGAGCTGGC 7784

Search completed: April 28, 2003, 15:56:39
Job time: 114.732 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.5
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: April 28, 2003, 13:51:29 ; Search time 2.33065 Seconds
(without alignments)
2236.936 Million cell updates/sec

Title: US-09-397-967-15
Perfect score: 90
Sequence: 1 AKLLPLDKYVREPG 17

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_plus.p2n.model -DEV=xlp
-O=/cgn2_1/USPRO.spool/US09397967/unat_26042003.182313.10750/app.query.fasta_1.1486
-DB=Issued_Patents.NA -GPMT=fastgap -SUFFIX=rml -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -STAMP=1 -END=1 -MATRIX=blosum62 -TRANS-human40.cdi
-LIST=45 -DOCCALITN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09397967 -CGN_1_1_59.etrunc_26042003.182313.10750 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG.SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: Issued_Patents.NA:*

1: /cgn2_6/pdata/2/1na/5A_COMB.seq:*
2: /cgn2_6/pdata/2/1na/5B_COMB.seq:*
3: /cgn2_6/pdata/2/1na/6A_COMB.seq:*
4: /cgn2_6/pdata/2/1na/6B_COMB.seq:*
5: /cgn2_6/pdata/2/1na/PCTUS_COMB.seq:*
6: /cgn2_6/pdata/2/1na/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	90	100.0	3807	1	US-08-357-598-1
2	90	100.0	3807	2	US-09-003-289-1
3	90	100.0	3807	5	PCT-US95-16435-1
4	83	92.2	4016	5	PCT-US95-08354A-1
5	64	71.1	3435	4	US-09-046-158A-21
6	64	71.1	3435	1	US-08-446-038B-2
7	64	71.1	3495	1	US-08-446-010B-2
8	64	71.1	3495	1	US-08-805-445-2
9	64	71.1	3495	2	US-08-064-067D-2
10	64	71.1	3495	2	US-09-066-208-2
11	64	71.1	3495	4	US-08-980-080-3
12	64	71.1	3629	1	US-08-097-997A-8

13	64	71.1	3629	3	US-08-665-574C-8	Sequence 8, Appl1
14	64	71.1	3629	4	US-08-946-994-8	Sequence 8, Appl1
15	64	71.1	4482	2	US-08-567-508C-1	Sequence 1, Appl1
16	64	71.1	4482	1	US-09-146-480-1	Sequence 1, Appl1
17	46	51.1	1863	2	US-08-455-073A-3	Sequence 3, Appl1
18	44	48.9	3147	4	US-09-101-886B-1	Sequence 3, Appl1
19	44	48.9	11703	4	US-09-101-886B-3	Sequence 3, Appl1
20	43	47.8	978	4	US-09-134-001C-2696	Sequence 2696, Ap
21	43	47.8	3561	3	US-08-097-997A-12	Sequence 12, Appl
22	43	47.8	3561	3	US-08-665-574C-12	Sequence 12, Appl
23	43	47.8	3561	4	US-08-946-994-12	Sequence 12, Appl
24	43	47.8	4052	2	US-08-833-226-1	Sequence 1, Appl1
25	43	47.8	5024	1	US-08-920-812-7	Sequence 7, Appl1
26	43	47.8	5024	1	US-08-920-812-7	Sequence 7, Appl1
27	43	47.8	5024	1	US-08-921-177-7	Sequence 7, Appl1
28	43	47.8	5024	1	US-08-362-577C-7	Sequence 7, Appl1
29	43	47.8	5024	2	US-08-920-828-7	Sequence 7, Appl1
30	42	46.7	2734	3	US-09-135-021-79	Sequence 79, Appl
31	42	46.7	2821	4	US-09-135-010A-115	Sequence 115, App
32	42	46.7	2821	4	US-09-597-735-115	Sequence 115, App
33	42	46.7	2821	4	US-09-597-732-115	Sequence 115, App
34	42	46.7	3181	3	US-09-135-021-1	Sequence 1, Appl1
35	42	46.7	3181	4	US-09-135-020-1	Sequence 1, Appl1
36	42	46.7	3181	4	US-09-135-010A-1	Sequence 1, Appl1
37	42	46.7	3181	4	US-09-444-871-1	Sequence 1, Appl1
38	42	46.7	3181	4	US-09-597-735-1	Sequence 1, Appl1
39	42	46.7	3181	4	US-09-444-295-1	Sequence 1, Appl1
40	42	46.7	3181	4	US-09-597-732-1	Sequence 1, Appl1
41	42	46.7	3181	3	US-09-135-021-5	Sequence 5, Appl1
42	41	45.6	998	4	US-09-540-014-26	Sequence 26, Appl
43	41	45.6	3355	2	US-08-933-821-3	Sequence 3, Appl1
44	41	45.6	3355	3	US-08-960-507-3	Sequence 3, Appl1
45	41	45.6	3355	4	US-09-136-828-3	Sequence 3, Appl1

ALIGNMENTS

RESULT 1
US-08-357-598-1
Sequence 1, Application US/08357598
Patent No. 5705625
GENERAL INFORMATION:
APPLICANT: Civin, Curt I.
TITLE OF INVENTION: NOVEL PROTEIN TYROSINE KINASE, JAK3
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/357,598
FILING DATE: 15-DEC-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Halle, Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07265/033001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3807 base pairs
TYPE: nucleic acid

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-357-598-1

Alignment Scores:

Pred. No.:	2.23e-07	Length:	3807
Score:	90.00	Matches:	17
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	1	Gaps:	0

US-09-397-967-15 (1-17) x US-08-357-598-1 (1-3807)

OY 1 AlalysleuLeuProLeuAspLysAspTyrTyryValValArgGluProGly 17
DB 3069 GCTAAGCTGCTGCCCTTGACAAAGACTACTACGTGGTCCGACGACGAGC 3119

RESULT 2

US-09-003-289-1
Sequence 1, Application US/09003289
Patent No. 5916792

GENERAL INFORMATION:

APPLICANT: Civin, Curt I.
APPLICANT: Small, Donald
TITLE OF INVENTION: NOVEL PROTEIN TYROSINE KINASE, JAK3
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla

STATE: CA
COUNTRY: USA

ZIP: 92037

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/003,289
FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/357,598
FILING DATE: 15-DEC-1994

ATTORNEY/AGENT INFORMATION:
NAME: Halle, Lisa A.

REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07265/033001

TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070

TELEFAX: 619/678-5099

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:
LENGTH: 3807 base pairs

TYPE: nucleic acid
STRANDEDNESS: linear
TOPOLOGY: linear

MOLECULE TYPE: DNA
US-09-003-289-1

Alignment Scores:

Pred. No.:	2.23e-07	Length:	3807
Score:	90.00	Matches:	17
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	2	Gaps:	0

US-09-397-967-15 (1-17) x US-09-003-289-1 (1-3807)

OY 1 AlalysleuLeuProLeuAspLysAspTyrTyryValValArgGluProGly 17
DB 3069 GCTAAGCTGCTGCCCTTGACAAAGACTACTACGTGGTCCGACGACGAGC 3119

RESULT 3

PCT-US95-16435-1
Sequence 1, Application PC/TUS9516435

GENERAL INFORMATION:

APPLICANT: The Johns Hopkins University School of Medicine
TITLE OF INVENTION: NOVEL PROTEIN TYROSINE KINASE, JAK3
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla

STATE: CA
COUNTRY: USA

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/16435
FILING DATE: 15-DEC-1995

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:
NAME: Halle, Lisa A.

REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07265/033001

TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5099

TELEFAX: 619/678-5070

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:
LENGTH: 3807 base pairs

TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: DNA
PCT-US95-16435-1

Alignment Scores:

Pred. No.:	2.23e-07	Length:	3807
Score:	90.00	Matches:	17
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	5	Gaps:	0

US-09-397-967-15 (1-17) x PCT-US95-16435-1 (1-3807)

OY 1 AlalysleuLeuProLeuAspLysAspTyrTyryValValArgGluProGly 17
DB 3069 GCTAAGCTGCTGCCCTTGACAAAGACTACTACGTGGTCCGACGACGAGC 3119

RESULT 4

PCT-US95-08354A-1
Sequence 1, Application PC/TUS9508354A

GENERAL INFORMATION:

APPLICANT: Temple University - Of The
Commonwealth System of Higher Education
TITLE OF INVENTION: JAK3 PROTEIN TYROSINE
KINASE AND DNA ENCODING THE SAME
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:

ADDRESSEE: Seidel, Gonda, Lavorgna
STREET: Suite 1800, Two Penn Center
CITY: Philadelphia

STATE: Pennsylvania
COUNTRY: U.S.A.

ZIP: 19102
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT-US95/08354A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/272,368
FILING DATE: 8 July 1994
ATTORNEY/AGENT INFORMATION:
NAME: Monaco, Daniel A.
REGISTRATION NUMBER: 30,480
REFERENCE/DOCKET NUMBER: 6056-203 PC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-8383
TELEFAX: (215) 568-5549
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4016 nucleotides
TYPE: nucleic acid
STRANDEDNESS: single stranded
TOPOLOGY: linear
PCT-US95-08354A-1

Alignment Scores:
Pred. No.: 4,97e-06 Length: 4016
Score: 83.00 Matches: 16
Percent Similarity: 94.12% Conservative: 0
Best Local Similarity: 94.12% Mismatches: 1
Query Match: 92.22% Indels: 0
DB: 5 Gaps: 0

US-09-397-967-15 (1-17) x PCT-US95-08354A-1 (1-4016)

QY 1 AAlaLysLeuLeuProLeuAspLysAspTyrTyrValValArgGluProGly 17
DB 3544 GCTAAGCTGCTGCCCTGGCAAGAGACTACTGCTGTCGCCAGACCTGGC 3594

RESULT 5
US-09-046-158A-21
Sequence 21, Application US/09046158A
Patent No. 6187552
GENERAL INFORMATION:
APPLICANT: Roberds, Steven L.
TITLE OF INVENTION: METHOD FOR IDENTIFYING INHIBITORS OF
TITLE OF INVENTION: JAK2/CYTOKINE RECEPTOR BINDING
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pharmacia & Upjohn Co., Intellectual Property
ADDRESS: Legal Services
STREET: 301 Henrietta Street
CITY: Kalamazoo
STATE: MI
COUNTRY: USA
ZIP: 49001
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/046,158A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Darnley Jr., James D.
REGISTRATION NUMBER: 33,673
TELECOMMUNICATION INFORMATION:

TELEPHONE: 616/833-2210
TELEFAX: 616/833-8897
TELEX: 224401
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 3435 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-046-158A-21

Alignment Scores:
Pred. No.: 0.0159 Length: 3435
Score: 64.00 Matches: 11
Percent Similarity: 87.50% Conservative: 3
Best Local Similarity: 68.75% Mismatches: 2
Query Match: 71.11% Indels: 0
DB: 4 Gaps: 0

US-09-397-967-15 (1-17) x US-09-046-158A-21 (1-3435)

QY 2 LysLeuLeuProLeuAspLysAspTyrTyrValValArgGluProGly 17
DB 2998 AAAGTCTTCCACACAGCAAGAAATCTATTAAGTAAGAAAGAACTGGT 3045

RESULT 6
US-08-446-038B-2
Sequence 2, Application US/08446038B
Patent No. 5658791
GENERAL INFORMATION:
APPLICANT: Wilks, Andrew F.; Ziemlecki, Andrew;
APPLICANT: Harpur, Alla
TITLE OF INVENTION: No. 5658791el Protein Tyrosine Kinase
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Felte & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/446,038B
FILING DATE: 19-MAY-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/064,067
FILING DATE: 30-Jun-1993
APPLICATION NUMBER: PCT/US91/08889
FILING DATE: 26-No. 5658791-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: Australian PK3594/90
FILING DATE: 28-No. 5658791-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: Australian 88229/91
FILING DATE: 27-No. 5658791-1991
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 5658791man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 5244
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-688-9200
TELEFAX: 212-838-3884
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 3495 base pairs
TYPE: nucleic acid

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: nucleic acid
US-08-446-038B-2

Alignment Scores:
Pred. No.: 0.0162 Length: 3495
Score: 64.00 Matches: 11
Percent Similarity: 87.50% Conservative: 3
Best Local Similarity: 68.75% Mismatches: 2
Query Match: 71.11% Indels: 0
DB: 1 Gaps: 0

US-09-397-967-15 (1-17) x US-08-446-038B-2 (1-3495)

Qy 2 LysleuProleuAspLysAspTyrTyrValValArgGluProgly 17
Db 2587 AAAGCTTGCCCGACAGACAAAGATCTACAAAGTAAAGAGCCAGGG 2634

RESULT 7
US-08-446-010B-2
Sequence 2, Application US/08446010B
Patent No. 5716818
GENERAL INFORMATION:
APPLICANT: Wilks, Andrew F.; Ziemiacki, Andrew;
APPLICANT: Harpur, Alisa
TITLE OF INVENTION: NO. 5716818el Protein Tyrosine Kinase
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Felfe & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/446,010B
FILING DATE: 19-May-1995
CLASSIFICATION: 433
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/446,038
FILING DATE: 19-May-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/064,067
FILING DATE: 30-Jun-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US91/08889
FILING DATE: 26-No. 5716818-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: Australian PK3594/90
FILING DATE: 28-No. 5716818-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: Australian 88229/91
FILING DATE: 27-No. 5716818-1991
ATTORNEY/AGENT INFORMATION:
NAME: Baer, Madeline F.
REGISTRATION NUMBER: 36,437
REFERENCE/DOCKET NUMBER: LUD 5244.3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-688-9200
TELEFAX: 212-838-3884
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 3495 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: nucleic acid

US-08-446-010B-2

Alignment Scores:
Pred. No.: 0.0162 Length: 3495
Score: 64.00 Matches: 11
Percent Similarity: 87.50% Conservative: 3
Best Local Similarity: 68.75% Mismatches: 2
Query Match: 71.11% Indels: 0
DB: 1 Gaps: 0

US-09-397-967-15 (1-17) x US-08-446-010B-2 (1-3495)

Qy 2 LysleuProleuAspLysAspTyrTyrValValArgGluProgly 17
Db 2587 AAAGCTTGCCCGACAGACAAAGATCTACAAAGTAAAGAGCCAGGG 2634

RESULT 8
US-08-805-445-2
Sequence 2, Application US/08805445
Patent No. 5821069
GENERAL INFORMATION:
APPLICANT: Wilks, Andrew F.; Ziemiacki, Andrew;
APPLICANT: Harpur, Alisa
TITLE OF INVENTION: NO. 5821069el Protein Tyrosine Kinase
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Felfe & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/805,445
FILING DATE: 25-FEB-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/446,038
FILING DATE: 19-MAY-1995
APPLICATION NUMBER: 08/064,067
FILING DATE: 30-Jun-1993
APPLICATION NUMBER: PCT/US91/08889
FILING DATE: 26-No. 5821069-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: Australian PK3594/90
FILING DATE: 28-No. 5821069-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: Australian 88229/91
FILING DATE: 27-No. 5821069-1991
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 5821069man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 5244
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-688-9200
TELEFAX: 212-838-3884
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 3495 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: nucleic acid

US-08-805-445-2

Alignment Scores:
Pred. No.: 0.0162 Length: 3495
Score: 64.00 Matches: 11

Percent Similarity: 87.508
Best Local Similarity: 68.758
Query Match: 71.118
DB: 1
Gaps: 0

US-09-397-967-15 (1-17) x US-08-805-445-2 (1-3495)

QY 2 LysleuleuProleuAspLysAspTyrTyValValArgLupProGly 17
DB 2587 AAGGCTTCCGCGAGCAAGAAATACTACAAAGTAAGAGACCGAGG 2634

RESULT 9

US-08-064-067D-2
Sequence 2, Application US/08064067D
Patent No. 5852184

GENERAL INFORMATION:

APPLICANT: Wilks, Andrew F.; Ziemlecki, Andrew;

APPLICANT: Harpur, Alisa

TITLE OF INVENTION: No. 5852184e1 Protein Tyrosine Kinase

NUMBER OF SEQUENCES: 23

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Felte & Lynch

STREET: 805 Third Avenue

CITY: New York City

STATE: New York

COUNTRY: USA

ZIP: 10022

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage

COMPUTER: IBM PS/2

OPERATING SYSTEM: PC-DOS

SOFTWARE: Wordperfect

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/064, 067D

FILING DATE: 30-Jun-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/US91/08889

FILING DATE: 26-No. 5852184-1991

PRIOR APPLICATION DATA:

APPLICATION NUMBER: Australian PK3594/90

FILING DATE: 28-No. 5852184-1990

PRIOR APPLICATION DATA:

APPLICATION NUMBER: Australian 88229/91

FILING DATE: 27-No. 5852184-1991

ATTORNEY/AGENT INFORMATION:

NAME: Hanson, No. 5852184man D.

REGISTRATION NUMBER: 30,946

REFERENCE/DOCKET NUMBER: LUD 5244

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-688-9200

TELEFAX: 212-838-3884

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 3495 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: nucleic acid

US-08-064-067D-2

RESULT 10

US-09-066-208-2
Sequence 2, Application US/09066208

Patent No. 5910426

GENERAL INFORMATION:

APPLICANT: Wilks, Andrew F.; Ziemlecki, Andrew;

APPLICANT: Harpur, Alisa

TITLE OF INVENTION: No. 5910426e1 Protein Tyrosine Kinase

NUMBER OF SEQUENCES: 23

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Felte & Lynch

STREET: 805 Third Avenue

CITY: New York City

STATE: New York

COUNTRY: USA

ZIP: 10022

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage

COMPUTER: IBM PS/2

OPERATING SYSTEM: PC-DOS

SOFTWARE: Wordperfect

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/066, 208

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/805, 445

FILING DATE: 25-FEB-1997

APPLICATION NUMBER: US 08/446, 038

FILING DATE: 19-MAY-1995

APPLICATION NUMBER: 08/064, 067

FILING DATE: 30-Jun-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/US91/08889

FILING DATE: 26-No. 5910426-1991

PRIOR APPLICATION DATA:

APPLICATION NUMBER: Australian PK3594/90

FILING DATE: 28-No. 5910426-1990

PRIOR APPLICATION DATA:

APPLICATION NUMBER: Australian 88229/91

FILING DATE: 27-No. 5910426-1991

ATTORNEY/AGENT INFORMATION:

NAME: Hanson, No. 5910426man D.

REGISTRATION NUMBER: 30,946

REFERENCE/DOCKET NUMBER: LUD 5244

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-688-9200

TELEFAX: 212-838-3884

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 3495 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: nucleic acid

US-09-066-208-2

RESULT 11

US-09-397-967-15 (1-17) x US-09-066-208-2 (1-3495)

QY 2 LysleuleuProleuAspLysAspTyrTyValValArgLupProGly 17
DB 2587 AAGGCTTCCGCGAGCAAGAAATACTACAAAGTAAGAGACCGAGG 2634

RESULT 11

US-08-980-080-3

Sequence 3, Application US/08980080
Patent No. 6312941
GENERAL INFORMATION:
APPLICANT: CARTER-SU, CHRISTIN
APPLICANT: RUI, LIANG-YOU
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING
TITLE OF INVENTION: SIGNALING PATHWAY AGONISTS AND ANTAGONISTS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: MEDLEN & CARROLL, LLP
STREET: 220 MONTGOMERY STREET, SUITE 2200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/980,080
FILING DATE: 26-NOV-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: CARROLL, PETER G.
REGISTRATION NUMBER: 32,837
REFERENCE/DOCKET NUMBER: UM-03036
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3495 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..2979
US-08-980-080-3

Alignment Scores:
Pred. No.: 0.0162 Length: 3495
Score: 64.00 Matches: 11
Percent Similarity: 87.50% Conservative: 3
Best Local Similarity: 68.75% Mismatches: 2
Query Match: 71.11% Indels: 0
Gaps: 0
DB: 4

US-09-397-967-15 (1-17) x US-08-980-080-3 (1-3495)

QY 2 LysleuLeuProLeuAspLysAspTyrTyrValValArgGluProGly 17
DB 2587 AAAGTCTTGCCGACGACAAAGAACTACTCAAAAGTAAGGACGAGG 2634
||||:||||| |||||:||||| |||||:||||| |||||:||||| |||||:|||||

RESULT 12
US-08-097-997A-8
Sequence 8, Application US/08097997A
Patent No. 5728536
GENERAL INFORMATION:
APPLICANT: Ihle, James N.
APPLICANT: Silvenoinen, Olli
APPLICANT: Wilthuhn, Bruce A.
APPLICANT: Quelle, Frederick W.
TITLE OF INVENTION: Jak Kinases and Regulation of Cytokine Signal
TITLE OF INVENTION: Transduction
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1100 New York Avenue, Suite 600
CITY: Washington

STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/097,997A
FILING DATE: 29-JULY-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Fox, Samuel L.
REGISTRATION NUMBER: 30,353
REFERENCE/DOCKET NUMBER: 0656.0370000/SLE/GKT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 3629 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 94..3480
US-08-097-997A-8

Alignment Scores:
Pred. No.: 0.0169 Length: 3629
Score: 64.00 Matches: 11
Percent Similarity: 87.50% Conservative: 3
Best Local Similarity: 68.75% Mismatches: 2
Query Match: 71.11% Indels: 0
Gaps: 0
DB: 1

US-09-397-967-15 (1-17) x US-08-097-997A-8 (1-3629)

QY 2 LysleuLeuProLeuAspLysAspTyrTyrValValArgGluProGly 17
DB 3088 AAAGTCTTGCCGACGACAAAGAACTACTCAAAAGTAAGGACGAGG 3135
||||:||||| |||||:||||| |||||:||||| |||||:||||| |||||:|||||

RESULT 13
US-08-665-574C-8
Sequence 8, Application US/08665574C
Patent No. 6136595
GENERAL INFORMATION:
APPLICANT: Ihle, James N.
APPLICANT: Silvenoinen, Olli
APPLICANT: Wilthuhn, Bruce A.
TITLE OF INVENTION: Signal Transduction
TITLE OF INVENTION: Regulation of Cytokine
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/665,574C
FILING DATE: 18-JUN-1996
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/282,012

FILED DATE: 29-JUL-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/097,997
FILING DATE: 29-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/118,968
FILING DATE: 09-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: Bugalsky, Lawrence B.
REGISTRATION NUMBER: 35,086
REFERENCE/DOCKET NUMBER: 0656,0370002/SLF/LBB
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2500
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 3629 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 94..3480
US-08-665-574C-8

Alignment Scores:
Pred. No.: 0.0169 Length: 3629
Score: 64.00 Matches: 11
Percent Similarity: 87.50% Conservative: 3
Best Local Similarity: 68.75% Mismatches: 2
Query Match: 71.11% Indels: 0
DB: 3 Gaps: 0

US-09-397-967-15 (1-17) x US-08-665-574C-8 (1-3629)

OY 2 LysleuleuProleuaspIysaspItyrTyValValAArgGluProGly 17
DB 3088 AAGGCTTGCCGACGACAAAGATCTACAAAGTAAAGAGCCAGG 3135
||||:||||| |||||:||||| |||||:|||||

RESULT 14
US-08-946-994-8
Sequence 8, Application US/08946994
Patent No. 6210634
GENERAL INFORMATION:
APPLICANT: Ihle, James N.
APPLICANT: Silvenoinen, Olli
APPLICANT: Wiltbuhn, Bruce A.
APPLICANT: Ouelle, Frederick W.
TITLE OF INVENTION: Jak kinases and Regulation of cytokine signal
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/946,994
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/665,574
FILING DATE: 18-JUN-1996
APPLICATION NUMBER: 08/282,012
FILING DATE: 29-JUL-1994
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/097,997
FILING DATE: 29-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/118,968
FILING DATE: 09-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: Fox, Samuel L.
REGISTRATION NUMBER: 30,353
REFERENCE/DOCKET NUMBER: 0656,0370002/SLF/GKT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 3629 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 94..3480
US-08-946-994-8

Alignment Scores:
Pred. No.: 0.0169 Length: 3629
Score: 64.00 Matches: 11
Percent Similarity: 87.50% Conservative: 3
Best Local Similarity: 68.75% Mismatches: 2
Query Match: 71.11% Indels: 0
DB: 4 Gaps: 0

US-09-397-967-15 (1-17) x US-08-946-994-8 (1-3629)

OY 2 LysleuleuProleuaspIysaspItyrTyValValAArgGluProGly 17
DB 3088 AAGGCTTGCCGACGACAAAGATCTACAAAGTAAAGAGCCAGG 3135
||||:||||| |||||:||||| |||||:|||||

RESULT 15
US-08-567-508C-1
Sequence 1, Application US/08567508C
Patent No. 5914393
GENERAL INFORMATION:
APPLICANT: Coleman, Roger
APPLICANT: Stuart, Susan G.
TITLE OF INVENTION: A NOVEL HUMAN JAK2 KINASE
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: US
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/567,508C
FILING DATE: 05-DEC-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-004905
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:
LENGTH: 4482 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
IMMEDIATE SOURCE:
LIBRARY: Placentia
CLONE: 179527
US-08-567-508C-1

Alignment Scores:
Pred. No.: 0.0217 Length: 4482
Score: 64.00 Matches: 11
Percent Similarity: 87.50% Conservative: 3
Best Local Similarity: 68.75% Mismatches: 2
Query Match: 71.11% Indels: 0
DB: 2 Gaps: 0

US-09-397-967-15 (1-17) x US-08-567-508C-1 (1-4482)

QY 2 LysleuLeuProLeuAspLysAspTyrTyrValValArgGIuProGly 17
||||:||||| | |||||:||||| | |||||:|||||
DB 3334 AAGCTCTGCCACACACAAAGATACCTATAAGTAAAGAAGAACTGGT 3381

Search completed: April 28, 2003, 16:57:51
Job time : 6.33065 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 28, 2003, 16:07:34 : Search time 4.67652 Seconds

(without alignments)
3955.543 Million cell updates/sec

Title: US-09-397-967-15
Perfect score: 90
Sequence: 1 AKLLPLDKDYVVRPG 17

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 709820 seqs, 544064369 residues

Total number of hits satisfying chosen parameters: 1419640

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n_model -DEV=xlp
-O=/cg2_1/USPRO.spool/US09397967/US09397967.unat_26042003_182314_10830/app.query.fasta_1.1486
-DB=Published Applications NA -OFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=100sum62
-TRANS=human40.cd1 -LIST=45 -DOCALIGN=200 -THR_SCORE=pcr -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=plc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09397967@CGN_1.1.202.@runat_26042003_182314_10830
-NCPU=6 -ICPU=3 -NO_XLPHY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT
-DISPATCH=100 -LONGLOG -DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10
-XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: Published Applications NA.*

1: /cg2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cg2_6/ptodata/1/pubpna/PC1_NEW_PUB.seq.*
3: /cg2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cg2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cg2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cg2_6/ptodata/1/pubpna/PC105_PUBCOMB.seq.*
7: /cg2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cg2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cg2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
10: /cg2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
11: /cg2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
12: /cg2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*
13: /cg2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
14: /cg2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	48	53.3	32190	9 US-10-091-504-2209 Sequence 2209, Ap
2	48	53.3	32190	10 US-09-764-869-2209 Sequence 2209, Ap
3	45	50.0	212	10 US-09-983-965-1232 Sequence 1232, Ap
4	45	50.0	481	9 US-10-092-154-1382 Sequence 1382, Ap

C	5	45	50.0	481	10	US-09-764-847-1382	Sequence 1382, Ap
C	6	45	50.0	861	10	US-09-770-445-589	Sequence 589, Ap
C	7	45	50.0	1152	9	US-09-938-842A-2486	Sequence 2486, Ap
C	8	45	50.0	22786	10	US-09-764-877-3727	Sequence 3727, Ap
C	9	44	48.9	405	10	US-09-960-352-1663	Sequence 1663, Ap
C	10	44	48.9	411	10	US-09-960-352-11844	Sequence 11844, A
C	11	44	48.9	430	9	US-09-736-457-48	Sequence 48, Ap
C	12	44	48.9	430	9	US-09-902-941-48	Sequence 48, Ap
C	13	44	48.9	430	9	US-09-849-626-48	Sequence 48, Ap
C	14	44	48.9	430	9	US-10-017-754-48	Sequence 48, Ap
C	15	44	48.9	430	10	US-09-998-598-1085	Sequence 1085, Ap
C	16	44	48.9	579	9	US-10-060-036-2884	Sequence 2884, Ap
C	17	44	48.9	1387	9	US-09-925-302-359	Sequence 359, Ap
C	18	44	48.9	1638	7	US-08-976-063C-9	Sequence 9, Ap
C	19	44	48.9	2013	9	US-09-938-842A-1519	Sequence 1519, Ap
C	20	44	48.9	3931	10	US-09-880-107-3365	Sequence 3365, Ap
C	21	44	48.9	32679	7	US-08-976-063C-1	Sequence 1, Ap
C	22	43	47.8	259	9	US-09-925-922-3	Sequence 3, Ap
C	23	43	47.8	365	10	US-09-878-574-2001	Sequence 2001, Ap
C	24	43	47.8	372	10	US-09-864-761-3611	Sequence 3611, Ap
C	25	43	47.8	410	10	US-09-864-761-32588	Sequence 32588, A
C	26	43	47.8	481	10	US-09-864-761-20379	Sequence 16094, A
C	27	43	47.8	654	10	US-09-864-761-16094	Sequence 6, Ap
C	28	43	47.8	1149	9	US-10-132-812-5	Sequence 5, Ap
C	29	43	47.8	1155	9	US-10-132-812-3	Sequence 3, Ap
C	30	43	47.8	1155	10	US-09-764-556-1	Sequence 1, Ap
C	31	43	47.8	1158	9	US-10-132-812-1	Sequence 1, Ap
C	32	43	47.8	1170	9	US-09-925-922-1	Sequence 1, Ap
C	33	43	47.8	1287	9	US-09-510-333-102	Sequence 102, Ap
C	34	43	47.8	1387	10	US-09-750-373-13	Sequence 13, Ap
C	35	43	47.8	2811	9	US-09-938-842A-2538	Sequence 2538, Ap
C	36	43	47.8	4846	7	US-08-781-986A-276	Sequence 276, Ap
C	37	43	47.8	42999	9	US-09-740-029-3	Sequence 3, Ap
C	38	43	47.8	44848	9	US-09-988-113-42	Sequence 42, Ap
C	39	43	47.8	44848	10	US-09-776-874A-42	Sequence 42, Ap
C	40	43	47.8	49744	10	US-09-927-091-4	Sequence 4, Ap
C	41	42.5	46.7	414	9	US-09-918-999-11687	Sequence 17687, A
C	42	42	46.7	470	9	US-09-918-999-22620	Sequence 23620, A
C	43	42	46.7	483	9	US-09-918-999-13854	Sequence 13854, A
C	44	42	46.7	2821	10	US-09-880-107-3358	Sequence 3358, Ap
C	45	42	46.7				

ALIGNMENTS

RESULT 1
US-10-091-504-2209 : Sequence 2209, Application US/10091504
: Publication No. US20030059908A1
: GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
: FILE REFERENCE: PC007C1
: CURRENT APPLICATION NUMBER: US/10/091,504
: CURRENT FILING DATE: 2002-03-07
: NUMBER OF SEQ ID NOS: 2442
: PRIOR Application removed - See File Wrapper or Palm
: SOFTWARE: Patencin Ver. 2.0
: SEQ ID NO 2209
: LENGTH: 32190
: TYPE: DNA
: ORGANISM: Homo sapiens
US-10-091-504-2209

Alignment Scores:
Pred. No.: 268
Score: 48.00
Percent Similarity: 75.00%
Best Local Similarity: 50.00%
Query Match: 53.33%
DB: 9
Gaps: 0
Length: 32190
Matches: 8
Conservative: 4
Mismatch: 4
Indels: 0

US-09-397-967-15 (1-17) x US-10-091-504-2209 (1-32190)

OY 2 LysleuleuproleuasplysAspTyrTyValValargluprogly 17
Db 19787 AGAATGACACTTGTGATGATGACTACTACTACTATTAACCTGGT 19834

RESULT 2
US-09-764-869-2209
; Sequence 2209, Application US/09764869
; Patent No. US20020061521A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007
; CURRENT APPLICATION NUMBER: US/09/764,869
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2442
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 2209
; LENGTH: 32190
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-869-2209

Alignment Scores:
Pred. No.: 268 Length: 32190
Score: 48.00 Matches: 8
Percent Similarity: 75.00% Conservative: 4
Best Local Similarity: 50.00% Mismatches: 4
Query Match: 53.33% Indels: 0
DB: 10 Gaps: 0

US-09-397-967-15 (1-17) x US-09-764-869-2209 (1-32190)
OY 2 LysleuleuproleuasplysAspTyrTyValValargluprogly 17
Db 19787 AGAATGACACTTGTGATGATGACTACTACTACTATTAACCTGGT 19834

RESULT 3
US-09-983-965-1232/C
; Sequence 1232, Application US/09983965
; Patent No. US20020137160A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 37-21(10297)C
; CURRENT APPLICATION NUMBER: US/09/983,965
; CURRENT FILING DATE: 2001-10-26
; PRIOR APPLICATION NUMBER: US 09/465,231
; PRIOR FILING DATE: 1999-12-15
; PRIOR APPLICATION NUMBER: US 60/113,678
; PRIOR FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 5912
; SEQ ID NO 1232
; LENGTH: 212
; TYPE: DNA
; ORGANISM: Bos taurus
; FEATURE:
; OTHER INFORMATION: Clone ID: 09-LIB2809-026-Q1-E1-C1
US-09-983-965-1232

Alignment Scores:
Pred. No.: 227 Length: 212
Score: 45.00 Matches: 8
Percent Similarity: 85.71% Conservative: 4
Best Local Similarity: 57.14% Mismatches: 2
Query Match: 50.00% Indels: 0
DB: 10 Gaps: 0

US-09-397-967-15 (1-17) x US-09-983-965-1232 (1-212)

OY 3 LeuleuproleuasplysAspTyrTyValValargluprogly 16
Db 189 CTACTGCTCTGACGAGAACAGTTCTTCTGTGATGACGACCG 148

RESULT 4
US-10-092-154-1382/C
; Sequence 1382, Application US/10092154
; Publication No. US20030054375A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC009C1
; CURRENT APPLICATION NUMBER: US/10/092,154
; CURRENT FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 2003
; Prior Application removed - See file Wrapper or Palm
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 1382
; LENGTH: 481
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-092-154-1382

Alignment Scores:
Pred. No.: 618 Length: 481
Score: 45.00 Matches: 7
Percent Similarity: 76.92% Conservative: 3
Best Local Similarity: 53.85% Mismatches: 3
Query Match: 50.00% Indels: 0
DB: 9 Gaps: 0

US-09-397-967-15 (1-17) x US-10-092-154-1382 (1-481)

OY 5 ProleuasplysAspTyrTyValValargluprogly 17
Db 248 CCTGTGACCGTGACAGATATGATATCAATATCCTCGA 210

RESULT 5
US-09-764-847-1382/C
; Sequence 1382, Application US/09764847
; Patent No. US20020132767A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC009
; CURRENT APPLICATION NUMBER: US/09/764,847
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2003
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 1382
; LENGTH: 481
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-847-1382

Alignment Scores:
Pred. No.: 618 Length: 481
Score: 45.00 Matches: 7
Percent Similarity: 76.92% Conservative: 3
Best Local Similarity: 53.85% Mismatches: 3
Query Match: 50.00% Indels: 0
DB: 10 Gaps: 0

US-09-397-967-15 (1-17) x US-09-764-847-1382 (1-481)

OY 5 ProleuasplysAspTyrTyValValargluprogly 17
Db 248 CCTGTGACCGTGACAGATATGATATCAATATCCTCGA 210

RESULT 6


```
US-09-770-445-589/C
; Sequence 589, Application US/09770445
; Patent No. US20020023281A1
; GENERAL INFORMATION:
; APPLICANT: Goriach, Jörn
; APPLICANT: An, Yong-Qiang
; APPLICANT: Hamilton, Carol M.
; APPLICANT: Price, Jennifer L.
; APPLICANT: Raines, Tracy M.
; APPLICANT: Yu, Yang
; APPLICANT: Rameaka, Joshua G.
; APPLICANT: Page, Amy
; APPLICANT: Mathew, Abraham V.
; APPLICANT: Ledford, Brooke L.
; APPLICANT: Woessner, Jeffrey P.
; APPLICANT: Haas, William David
; APPLICANT: Garcia, Carlos A.
; APPLICANT: Krieger, MaJa
; APPLICANT: Slader, Ted
; APPLICANT: Davis, Keith R.
; APPLICANT: Allen, Keith
; APPLICANT: Hoffman, Neil
; APPLICANT: Hurban, Patrick
; TITLE OF INVENTION: Expressed Sequences of Arabidopsis
; FILE REFERENCE: 2023US (PARA-012PRV)
; CURRENT APPLICATION NUMBER: US/09/770,445.
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: US 60/178,472
; PRIOR FILING DATE: 2000-01-27
; NUMBER OF SEQ ID NOS: 999
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 589
; LENGTH: 861
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-770-445-589

Alignment Scores:
Pred. No.: 12.6 Length: 861
Score: 45.00 Matches: 9
Percent Similarity: 78.57% Conservative: 2
Best Local Similarity: 64.29% Mismatches: 3
Query Match: 50.00% Indels: 0
DB: 10 Gaps: 0

US-09-397-967-15 (1-17) x US-09-770-445-589 (1-861)
QY 4 LeuProLeuAspLysAspTyrTyrValValArgGluProGly 17
DB 524 CTTGACCTTGATGAGATGTTATGTTGTGACCAAGCCAGGT 483

RESULT 7
US-09-938-842A-2486
; Sequence 2486, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Krepes, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SCRIPT300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
```

```
; SEQ ID NO 2486
; LENGTH: 1152
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-2486

Alignment Scores:
Pred. No.: 17.9 Length: 1152
Score: 45.00 Matches: 9
Percent Similarity: 78.57% Conservative: 2
Best Local Similarity: 64.29% Mismatches: 3
Query Match: 50.00% Indels: 0
DB: 9 Gaps: 0

US-09-397-967-15 (1-17) x US-09-938-842A-2486 (1-1152)
QY 4 LeuProLeuAspLysAspTyrTyrValValArgGluProGly 17
DB 952 CTTGACCTTGATGAGATGTTATGTTGTGACCAAGCCAGGT 993

RESULT 8
US-09-764-877-3727/C
; Sequence 3727, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3727
; LENGTH: 22786
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-877-3727

Alignment Scores:
Pred. No.: 683 Length: 22786
Score: 45.00 Matches: 7
Percent Similarity: 84.62% Conservative: 4
Best Local Similarity: 53.85% Mismatches: 2
Query Match: 50.00% Indels: 0
DB: 10 Gaps: 0

US-09-397-967-15 (1-17) x US-09-764-877-3727 (1-22786)
QY 4 LeuProLeuAspLysAspTyrTyrValValArgGluPro 16
DB 6948 ATCCCCCGAATGAGACTATATACATTGTGTAGAAACCT 6910

RESULT 9
US-09-960-352-1663
; Sequence 1663, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nenping
; APPLICANT: Byatt, John C.
; APPLICANT: Mathalagan, Nagappan
; TITLE OF INVENTION: NOCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 1663
; LENGTH: 405
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 08-LIB188-026-Q1-E1-B7
```

US-09-960-352-1663

Alignment Scores:

Pred. No.:	7.88	Length:	405
Score:	44.00	Matches:	7
Percent Similarity:	78.57%	Conservative:	4
Best Local Similarity:	50.00%	Mismatches:	3
Query Match:	48.89%	Indels:	0
DB:	10	Gaps:	0

US-09-397-967-15 (1-17) x US-09-960-352-1663 (1-405)

OY 4 LeuProLeuaspIysAspTyTyValValargIuproGly 17

Db 229 CTGCCTGTCCACCAACAGCTTACTCTCTGAAGACCCCGGT 270

RESULT 10

US-09-960-352-11844/C
Sequence 11844, Application US/09960352
Patent No. US20020137139A1

GENERAL INFORMATION:

APPLICANT: Warren, Wesley C.
APPLICANT: Tao, Nengding
APPLICANT: Byatt, John C.
APPLICANT: Mathialagan, Nagappan
TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
FILE REFERENCE: 16511.006/37-21(10298)C
CURRENT APPLICATION NUMBER: US/09/960.352
CURRENT FILING DATE: 2001-09-24
NUMBER OF SEQ ID NOS: 15112
SEQ ID NO 11844
LENGTH: 411
TYPE: DNA
ORGANISM: Bos taurus
OTHER INFORMATION: Clone ID: 51-LIB188-010-01-E1-E4
US-09-960-352-11844

Alignment Scores:

Pred. No.:	8.02	Length:	411
Score:	44.00	Matches:	7
Percent Similarity:	78.57%	Conservative:	4
Best Local Similarity:	50.00%	Mismatches:	3
Query Match:	48.89%	Indels:	0
DB:	10	Gaps:	0

US-09-397-967-15 (1-17) x US-09-960-352-11844 (1-411)

OY 4 LeuProLeuaspIysAspTyTyValValargIuproGly 17

Db 211 CTGCCTGTCCACCAACAGCTTACTCTCTGAAGACCCCGGT 170

RESULT 11

US-09-736-457-48
Sequence 48, Application US/09736457
Patent No. US20020168637A1

GENERAL INFORMATION:

APPLICANT: Wang, Tongtong
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Lodes, Michael A.
APPLICANT: Fanger, Gary
APPLICANT: Vedivick, Tom
APPLICANT: Carter, Darriek
APPLICANT: Retter, Marc
APPLICANT: Mannion, Jane
APPLICANT: Fan, Liqun
APPLICANT: Wang, Aijun
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.478C15
CURRENT APPLICATION NUMBER: US/09/736.457
CURRENT FILING DATE: 2000-12-13
NUMBER OF SEQ ID NOS: 1864

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 48
LENGTH: 430
TYPE: DNA
ORGANISM: Homo sapien
US-09-736-457-48

Alignment Scores:

Pred. No.:	8.48	Length:	430
Score:	44.00	Matches:	11
Percent Similarity:	68.42%	Conservative:	2
Best Local Similarity:	57.89%	Mismatches:	4
Query Match:	48.89%	Indels:	2
DB:	9	Gaps:	1

US-09-397-967-15 (1-17) x US-09-736-457-48 (1-430)

OY 1 AlalysLeuPro-----LeuaspIysAspTyTyValValargIuproGly 17

Db 159 GCAAGCTTCCTCCAGTATTAGACCGATCTCTATTATTATTAATGCTGT 215

RESULT 12

US-09-902-941-48
Sequence 48, Application US/09902941
Patent No. US20020172952A1

GENERAL INFORMATION:

APPLICANT: Henderson, Robert A.
APPLICANT: Wang, Tongtong
APPLICANT: Matanabe, Yoshihiro
APPLICANT: Johnson, Jeffrey C.
APPLICANT: Retter, Marc W.
APPLICANT: Marnerakis, Margarita
APPLICANT: Carter, Darriek
APPLICANT: Fanger, Gary R.
APPLICANT: Vedivick, Thomas S.
APPLICANT: Bangur, Chaitanya S.
APPLICANT: McNabb, Andria
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
FILE REFERENCE: 210121.478C17
CURRENT APPLICATION NUMBER: US/09/902.941
CURRENT FILING DATE: 2001-07-10
NUMBER OF SEQ ID NOS: 2002
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 48
LENGTH: 430
TYPE: DNA
ORGANISM: Homo sapiens
US-09-902-941-48

Alignment Scores:

Pred. No.:	8.48	Length:	430
Score:	44.00	Matches:	11
Percent Similarity:	68.42%	Conservative:	2
Best Local Similarity:	57.89%	Mismatches:	4
Query Match:	48.89%	Indels:	2
DB:	9	Gaps:	1

US-09-397-967-15 (1-17) x US-09-902-941-48 (1-430)

OY 1 AlalysLeuPro-----LeuaspIysAspTyTyValValargIuproGly 17

Db 159 GCAAGCTTCCTCCAGTATTAGACCGATCTCTATTATTATTAATGCTGT 215

RESULT 13

US-09-849-626-48
Sequence 48, Application US/09849626
Publication No. US20020197659A1

GENERAL INFORMATION:

APPLICANT: Bangur, Chaitanya
APPLICANT: Fanger, Gary
APPLICANT: Wang, Aijun
APPLICANT: Wang, Tongtong

```

; APPLICANT: Switzer, Anne
; APPLICANT: McNeill, Patricia
; APPLICANT: Clapper, Jonathan
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C16
; CURRENT APPLICATION NUMBER: US/09/849,626
; NUMBER OF SEQ ID NOS: 1926
; SOFTWARE: FastSeq for Windows Version 3.0
; LENGTH: 430
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-849-626-48

Alignment Scores:
Pred. No.: 8.48 Length: 430
Score: 44.00 Matches: 11
Percent Similarity: 68.42% Conservative: 2
Best Local Similarity: 57.89% Mismatches: 4
Query Match: 48.89% Indels: 2
DB: 9 Gaps: 1

US-09-397-967-15 (1-17) x US-09-849-626-48 (1-430)

OY 1 AlalysleuleuPro-----leuaspysaspyrtyryvalvalargluProgly 17
DB 159 GCAAGCTTCTCCCTCACTATTAGACCGAGATCTCTATATTGTTAATGATGCT 215

RESULT 14
US-10-017-754-48
; Sequence 48, Application US/10017754
; Publication No. US20030054363A1
; GENERAL INFORMATION:
; APPLICANT: Henderson, Robert A.
; APPLICANT: Mang, Tonglong
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Johnson, Jeffrey C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Marnerakis, Margarita
; APPLICANT: Carter, Darick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: McNabb, Andria
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.478C18
; CURRENT APPLICATION NUMBER: US/10/017,754
; CURRENT FILING DATE: 2001-10-29
; NUMBER OF SEQ ID NOS: 2004
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 48
; LENGTH: 430
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-017-754-48

Alignment Scores:
Pred. No.: 8.48 Length: 430
Score: 44.00 Matches: 11
Percent Similarity: 68.42% Conservative: 2
Best Local Similarity: 57.89% Mismatches: 4
Query Match: 48.89% Indels: 2
DB: 9 Gaps: 1

US-09-397-967-15 (1-17) x US-10-017-754-48 (1-430)

OY 1 AlalysleuleuPro-----leuaspysaspyrtyryvalvalargluProgly 17
DB 159 GCAAGCTTCTCCCTCACTATTAGACCGAGATCTCTATATTGTTAATGATGCT 215

```

```

RESULT 15
US-09-998-598-1085
; Sequence 1085, Application US/09998598
; Patent No. US20020150922A1
; GENERAL INFORMATION:
; APPLICANT: Stoik, John A.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Cheanault, Ruth A.
; APPLICANT: Meagher, Madelein Joy
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.561
; CURRENT APPLICATION NUMBER: US/09/998,598
; CURRENT FILING DATE: 2001-11-16
; NUMBER OF SEQ ID NOS: 2606
; SOFTWARE: Corixa Invention Disclosure Database
; SEQ ID NO 1085
; LENGTH: 430
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-998-598-1085

Alignment Scores:
Pred. No.: 8.48 Length: 430
Score: 44.00 Matches: 11
Percent Similarity: 68.42% Conservative: 2
Best Local Similarity: 57.89% Mismatches: 4
Query Match: 48.89% Indels: 2
DB: 10 Gaps: 1

US-09-397-967-15 (1-17) x US-09-998-598-1085 (1-430)

OY 1 AlalysleuleuPro-----leuaspysaspyrtyryvalvalargluProgly 17
DB 159 GCAAGCTTCTCCCTCACTATTAGACCGAGATCTCTATATTGTTAATGATGCT 215

Search completed: April 28, 2003, 18:56:43
Job time : 10.6765 secs

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 28, 2003, 13:05:49 ; Search time 9.06362 Seconds
(without alignments)
4223.916 Million cell updates/sec

Title: US-09-397-967-15
Sequence: 1 AKLPLDKDYVYVREPG 17

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+g2n.model -DEV=xlp
-O=/cgn2.1/USPTO-spool/US09397967/runat_26042003_182312_10702/app.query.fasta_1.1486
-DB=N_Geneseq_101002 -OFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITS=bits -STRAP=1 -END=1 -MATRIX=plosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pcr -NOR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09397967.ecgn_1.1_418_etrunat_26042003_182312_10702 -NCPU=6 -ICPU=3
-NO_XLPHY -NO_MMAP -LARGEQUERY -NEG.SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV.TIMEOUT=120 -WARN.TIMEOUT=0 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FCGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: N_Geneseq_101002:*

1:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	90	100.0	165	16	AA087512	Tyrosine kinase pe
2	90	100.0	778	22	AA087512	Human Janus kinase
3	90	100.0	3807	17	AA087512	Protein tyrosine k
4	83	92.2	4016	17	AA087512	Mouse JAK3 protein
5	64	71.1	3435	19	AA087512	JAK2 protein encod
6	64	71.1	3473	13	AA087512	JAK2 encoding DNA
7	64	71.1	3495	24	AA087512	Human JAK2 kinase
8	64	71.1	3629	16	AA087512	Human JAK2 kinase
9	64	71.1	3629	21	AA087512	Human JAK2 kinase
10	64	71.1	3629	22	AA087512	Human JAK2 kinase
11	64	71.1	4482	20	AA087512	Human JAK2 kinase
12	64	71.1	4482	21	AA087512	Human JAK2 kinase
13	64	71.1	5117	24	AA087512	Human JAK2 kinase
14	51.5	57.2	1289	23	ABL21107	Drosophila melanog
15	51.5	57.2	2743	23	ABL21107	Drosophila melanog
16	51.5	57.2	3419	23	ABL21107	Drosophila melanog
17	51.5	57.2	3530	23	ABL21107	Drosophila melanog
18	49	54.4	835	24	ABN85718	Human membrane-ass
19	48	53.3	32190	22	AA087512	Human cardiovascu
20	46	51.1	668	23	AA087512	Drosophila melanog
21	46	51.1	687	23	ABN85718	Human cardiovascu
22	46	51.1	1863	18	AA087512	Drosophila melanog
23	46	51.1	2065	23	AA087512	Human FRK1 gene
24	46	51.1	2144	23	AA087512	DNA encoding novel
25	46	51.1	2144	23	AA087512	DNA encoding novel
26	45	50.0	481	23	ABK42495	Drosophila melanog
27	45	50.0	861	24	ABN98821	Human cardiovascu
28	45	50.0	1152	21	AA087512	Drosophila melanog
29	45	50.0	1360	22	AA087512	Human cardiovascu
30	45	50.0	1375	21	AA087512	Human cardiovascu
31	45	50.0	1375	21	AA087512	Human cardiovascu
32	45	50.0	22786	22	AA087512	Human cardiovascu
33	45	50.0	24887	24	ABK14141	Human cardiovascu
34	44	48.9	430	22	AA087512	Human cardiovascu
35	44	48.9	430	22	AA087512	Human cardiovascu
36	44	48.9	1143	24	ABK8041	Human cardiovascu
37	44	48.9	1165	23	AA087512	Human cardiovascu
38	44	48.9	1266	24	ABN85718	Human cardiovascu
39	44	48.9	1387	21	AA087512	Human cardiovascu
40	44	48.9	1387	21	AA087512	Human cardiovascu
41	44	48.9	2013	21	AA087512	Human cardiovascu
42	44	48.9	2068	21	AA087512	Human cardiovascu
43	44	48.9	2070	21	AA087512	Human cardiovascu
44	44	48.9	3147	18	AA087512	Human cardiovascu
45	44	48.9	3147	18	AA087512	Human cardiovascu

ALIGNMENTS

RESULT 1
AA087512
AA087512 standard; cDNA to mRNA; 165 BP.
AA087512:
22-NOV-1995 (first entry)
Tyrosine kinase peptide coding sequence.
Tyrosine kinase peptide: UT-7 cell; primer: polymerase chain reaction;
PCR: cell differentiation; antisense drug; ds.
Homo sapiens.
JP07059569-A.
07-MAR-1995.

```

XX PF 25-AUG-1993; 93JP-0210404.
XX PR 25-AUG-1993; 93JP-0210404.
XX PA (ASAH ) ASAH KASEI KOGYO KK.
XX DR WPI: 1995-135894/18.
XX DR P-PSDB; AAR71394.
XX PT New DNA sequence encoding a tyrosine kinase peptide - useful for
XX evaluation and control of cell differentiation.
XX PS Claim 1: Page 5; 5pp; Japanese.
XX CC This sequence encodes a tyrosine kinase peptide which is derived
XX from UT-7 cells. This sequence was isolated using the primer
XX CC sequences given in AA087510-11. The tyrosine kinase peptide may be
XX CC used for the evaluation and control of cell differentiation and is
XX CC also useful for the development of drugs and antisense drugs.
XX SO Sequence 165 BP; 36 A; 56 C; 39 G; 34 T; 0 other:

Alignment Scores:
Pred. No.: 2,29e-08 Length: 165
Score: 90.00 Matches: 17
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

US-09-397-967-15-(1-17) x AA087512 (1-165)
OY 1 AlalysleuProleuAspLysAspTyrTyrValAlaArgLupProGly 17
Db 52 GCTAAGCTGCTGCGCTTGACAAAGACTACTACTGCTGCTGCGGAGCCAGGC 102

RESULT 2
AAS10809
ID AAS10809 standard; cDNA; 778 BP.
XX AC AAS10809;
XX DT 24-OCT-2001 (first entry)
XX DE Human Janus kinase 3 (JAK3) partial cDNA sequence, 3-2 primer 1.
XX KW Human; Janus kinase 3; JAK3; ss; JAK/STAT inhibitor; 3-2 primer 1;
XX KW signal transducer and activator of transcription; osteoarthritis;
XX KW degenerative joint disease; rheumatoid arthritis; leprosy; asthma;
XX KW cancer; tumour; leukemia.
XX OS Homo sapiens.
XX PN WO200152892-A2.
XX PD 26-JUL-2001.
XX PF 22-JAN-2001; 2001WO-US02033.
XX PR 24-JAN-2000; 2000US-0177872.
XX PR 28-NOV-2000; 2000US-0723490.
XX PA (GENZ ) GENZYME CORP.
XX PI Vastios G;
XX DR WPI: 2001-465338/50.
XX PT Use of inhibitors of Janus kinase/signal transducers and activators of
XX PT transcription for inhibiting onset and progression of degenerative
XX PT joint diseases or disorders such as osteoarthritis, rheumatoid
XX PT arthritis

```

```

XX PS Example 4; Fig 3; 55pp; English.
XX CC The sequence is a partial sequence, designated "clone 3-2 primer 1",
XX CC obtained from a chondrocyte cDNA encoding human Janus kinase
XX CC 3 (JAK3). The invention relates to the use of JAK/STAT (Janus
XX CC kinase/signal transducer and activator of transcription)
XX CC inhibitors other than debromylenaldisine (DBM) and
XX CC mylenaldisine (H) for inhibiting the progression or the likelihood of
XX CC developing diseases involving cartilage degradation, and for regulating
XX CC the expression of pro-inflammatory agents or cytokines in a chondrocyte,
XX CC and cartilage degrading enzymes in a cell. A JAK/STAT inhibitor of the
XX CC invention is useful for inhibiting progression or likelihood of
XX CC developing osteoarthritis or rheumatoid arthritis. The inhibitor is also
XX CC useful for treating other JAK/STAT-mediated diseases or disorders,
XX CC including T cell-mediated disorders, mast cell-mediated disorders,
XX CC type 2 (cytokine hypersensitivity) disorders, B cell lymphoma, and
XX CC myeloid diseases. T cell-mediated disorders include human T cell
XX CC leukemia/lymphoma virus (HTLV)-1, Sdzory's syndrome, c-abl
XX CC transformation, natural killer-like T cell lymphomas, (NK-like tumours)
XX CC and graft-vs-host disease; cytokine hypersensitivity disorders include
XX CC leishmaniasis, leprosy, allergy and viral infections; mast cell-mediated
XX CC disorders include hay fever, asthma, hives and anaphylaxis; and
XX CC leukaemias and lymphomas include acute lymphocytic and lymphoblastic
XX CC leukaemias, B cell lymphomas and leukaemias of myeloid origin. DBM and H
XX CC are useful as therapeutic agents in cancers in which JAK3 plays a role
XX CC in the initiation or progression of tumorigenesis.
XX SO Sequence 778 BP; 131 A; 264 C; 220 G; 163 T; 0 other:

Alignment Scores:
Pred. No.: 1.41e-07 Length: 778
Score: 90.00 Matches: 17
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

US-09-397-967-15 (1-17) x AAS10809 (1-778)
OY 1 AlalysleuProleuAspLysAspTyrTyrValAlaArgLupProGly 17
Db 187 GCTAAGCTGCTGCGCTTGACAAAGACTACTACTGCTGCTGCGGAGCCAGGC 237

RESULT 3
AAT30862
ID AAT30862 standard; DNA; 3807 BP.
XX AC AAT30862;
XX DT 13-SEP-1996 (first entry)
XX DE Protein tyrosine kinase JAK3 cDNA.
XX KW JAK3; protein tyrosine kinase; cell proliferation; differentiation;
XX KW signal transduction; leukaemia; aplastic anaemia; myelodysplasia;
XX KW polythemia vera; thrombocytosis; gene therapy; diagnosis; ss.
XX OS Homo sapiens.
XX PN WO9618639-A1.
XX PD 20-JUN-1996.
XX PF 15-DEC-1995; 95WO-US16435.

Location/Qualifiers
Key 168..3359
FT CDS /*tag= a
FT polyA_signal 3692..3696
FT polyA_signal /*tag= b
FT polyA_signal 3793..3797
FT /*tag= c

```

```

XX 15-DEC-1994; 94US-0357598.
PR (0300) UNIV JOHNS HOPKINS SCHOOL MED.
PA C14v1n C1, Safford MG, Small D;
XX WPI: 1996-300568/30.
DR P-PSDB: AAR96037.
XX
XX protein tyrosine kinase, JAK3, protein and nucleic acid - used in
PT the gene therapy of cellular proliferative diseases, e.g. leukaemia,
PR aplastic anaemia etc.
XX
XX Claim 4: Page 43-47; 97pp; English.
XX
XX A cDNA clone (AAT30862) codes for JAK3 (AAR96037), a new member of the
CC JAK family of non-receptor protein tyrosine kinases, that probably
CC plays a role in growth factor modulated differentiation,
CC proliferation and survival of haematopoietic stem/progenitor cells.
CC It was derived from CD34+ mRNA obtd. from normal human bone marrow
CC by PCR amplification using primers (see also AAT30863-64) based
CC on highly conserved motifs from protein tyrosine kinase catalytic
CC domains. The JAK3 coding sequence can be incorporated into a
CC vector and used for prodn. of recombinant JAK3. It can also be
CC used in gene therapy protocols for leukaemia, myelodysplasia,
CC polthemia vera, thrombocytosis and aplastic anaemia, or to
CC stimulate haematopoietic cell proliferation. The gene was
CC localised to chromosome 19, band p12-13.1.
XX
XX Sequence 3807 BP; 716 A; 1219 C; 1099 G; 773 T; 0 other;
SQ
Alignment Scores:
Pred. No.: 9,066-07 Length: 3807
Score: 90.00 Matches: 17
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 17 Gaps: 0
US-09-397-967-15 (1-17) x AAT30862 (1-3807)
OY 1 AlAlysLeuLeuPProLeuAspTyrTyrValValArgGluProGly 17
DB 3069 GCTAAAGCTGCTGCCCTGACAAAGACTACTACGTGTCGCGAGCCAGCC 3119
RESULT 4
AAT11083
ID AAT11083 standard; cDNA; 4016 BP.
XX
XX AAT11083;
AC
XX
XX 08-APR-1996 (first entry)
DT
XX
XX Mouse JAK3 protein-tyrosine-kinase gene.
DE
XX
XX Mouse; Janus kinase; JAK3; protein-tyrosine-kinase; cytokine;
KW signal transduction; 32Dc13; reverse transcription; PCR;
KW polymerase chain reaction; primer; interleukin-3;
KW granulocyte-macrophage colony stimulating factor; cDNA probe;
KW granulocyte colony stimulating factor; Escherichia coli;
KW haematopoietic cell; differentiation; leukaemia; therapy; ss.
XX
XX Mus musculus.
OS
XX
XX Key Location/Qualifiers
FH CDS 7..3906
FT /*tag= a
FT /product= JAK3 protein-tyrosine-kinase
FT /note= "EC-2.7.1.112"
XX
XX MO9601838-A1.
XX

```

```

PD 25-JAN-1996.
XX
XX 28-JUN-1995; 95MO-US08354.
PF
XX
XX 08-JUL-1994; 94US-0272368.
PR
XX
XX (UTEM) UNIV TEMPLE.
PA
XX
XX Rane SG, Reddy PE;
PI
XX
XX WPI: 1996-097584/10.
DR
XX
XX P-PSDB: AAR88560.
XX
XX JAK3 protein tyrosine kinase and DNA encoding it - useful to induce
PT differentiation of haematopoietic cells to treat or control
PR leukaemia
XX
XX Claim 4: Page 26-28; 49pp; English.
XX
XX The sequence encodes a novel protein-tyrosine-kinase of the Janus
CC kinase (JAK) family (involved in cytokine signal transduction),
CC designated JAK3. The sequence has been isolated from mouse 32Dc13
CC cells growing in interleukin-3 or granulocyte-macrophage colony
CC stimulating factor by reverse transcription-polymerase chain
CC reaction amplification, using primers J3 and J4 (AAT11084-T11085),
CC based on conserved sequences in JAK JH2 and JH1 domains. A 950-bp
CC cDNA probe is isolated, which is then used to screen a granulocyte
CC colony stimulating factor-treated 32Dc13 cDNA library in phage
CC lambda-gt11 in Escherichia coli to isolate the full-length gene.
CC The JAK3 protein is useful for stimulating the differentiation of
CC haematopoietic cells in therapy or control of leukaemia.
XX
XX Sequence 4016 BP; 742 A; 1203 C; 1235 G; 836 T; 0 other;
SQ
Alignment Scores:
Pred. No.: 2,056-05 Length: 4016
Score: 83.00 Matches: 16
Percent Similarity: 94.12% Conservative: 0
Best Local Similarity: 94.12% Mismatches: 1
Query Match: 92.22% Indels: 0
DB: 17 Gaps: 0
US-09-397-967-15 (1-17) x AAT11083 (1-4016)
OY 1 AlAlysLeuLeuPProLeuAspTyrTyrValValArgGluProGly 17
DB 3544 GCTAAAGCTGCTGCCCTGCGAAGGACTACTACGTGTCGCGAGCCCTGCG 3594
RESULT 5
AAV61801
ID AAV61801 standard; cDNA; 3435 BP.
XX
XX AAV61801;
AC
XX
XX 20-JAN-1999 (first entry)
DT
XX
XX JAK2 protein encoding cDNA.
DE
XX
XX JAK2 protein; cytoplasmic domain; beta subunit; screening; asthma;
KW interleukin; granulocyte macrophage-colony stimulating factor; GM-CSF;
KW IL-3; IL-5; human; ss.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH CDS 10..3429
FT /*tag= a
FT /product= "JAK2 protein"
XX
XX MO9843087-A1.
XX
XX 01-OCT-1998.
XX

```

```

PF 23-MAR-1998; 98WO-US05387.
XX
XX 24-MAR-1997; 97US-0041511.
XX
XX (PHAA ) PHARMACIA & UPJOHN CO.
XX
XX Kaytes PS, Roberds SL;
XX
XX WPI; 1998-532151/45.
XX
XX P-PSDB; AAW76425.
XX
XX Screening for compounds useful for preventing or treating asthma -
XX by determining if compounds inhibit binding of the JAK2 protein to
XX e.g. IL-3, IL-5 or GM-CSF
XX
XX Disclosure; Fig 22A-C; 112pp; English.
XX
XX This cDNA encodes the full-length JAK2 protein. The N-terminal fragment
XX of JAK2 protein can be used in the method of the invention of screening
XX for compounds useful for treating or preventing asthma. The method
XX comprises contacting a molecule comprising at least the N-terminal 294
XX amino acid residues of the JAK2 protein, with another molecule comprising
XX at least 13 membrane-proximal cytoplasmic amino acids of interleukin
XX (IL)-3, IL-5 or granulocyte macrophage-colony stimulating factor (GM-CSF)
XX proteins in the presence of the candidate compound, and determining
XX whether the first and the second molecules form a complex. If the
XX compound inhibits complex formation, it can be used to treat asthma.
XX
XX Sequence 3435 BP; 1157 A; 598 C; 733 G; 947 T; 0 other;
XX
XX Alignment Scores:
XX Pred. No.: 0.0682 Length: 3435
XX Score: 64.00 Matches: 11
XX Percent Similarity: 87.50% Conservative: 3
XX Best Local Similarity: 68.75% Mismatches: 2
XX Query Match: 71.11% Indels: 0
XX DB: Gaps: 0
XX
XX US-09-397-967-15 (1-17) x AAV61801 (1-3435)
XX
XX QY 2 LysleuleuProLeuAspLysAspTyrTyryValValArgLupProGly 17
XX |||||:||||| |||||:||||| |||||:|||||
XX Db 2998 AAGGCTTGCCACAGACAAAGAAATACATATAAGTAAAGAACCTGGT 3045
XX
XX RESULT 6
XX AAQ25307
XX ID AAQ25307 standard; DNA; 3473 BP.
XX
XX AC AAQ25307;
XX
XX DT 18-FEB-1999 (first entry)
XX
XX DE JAK2 encoding DNA.
XX
XX DE JAK2 encoding DNA.
XX
XX KW Phosphorylation; JAK1; JAK2; protein tyrosine kinase; human;
XX catalytic domain; SH2 domain; growth factor receptor; PTK; murine; ss.
XX
XX OS Mus musculus.
XX
XX Key Location/Qualifiers
XX FT cds 1..2961
XX FT /*tag= a
XX FT /product= JAK2
XX
XX WO9210519-A.
XX
XX PD 25-JUN-1992.
XX
XX XX 26-NOV-1991; 91WO-US08889.
XX
XX PF 26-NOV-1991; 91WO-US08889.
XX
XX PR 28-NOV-1990; 90AU-0003594.
XX
XX (LUDW-) LUDWIG INST CANCER RES.
XX

```

```

XX
XX Harpur A, Wilks AF, Ziemiecki A;
XX
XX WPI; 1992-234591/28.
XX
XX P-PSDB; AAR25141.
XX
XX Novel protein tyrosine kinase mol. - comprises multiple catalytic
XX domains but no SH2 domain and is for phosphorylation of proteins
XX
XX Claim 10; Fig 8; 50pp; English.
XX
XX This sequence encodes the murine protein tyrosine kinase JAK2 (from
XX Janus kinase). Northern analysis of JAK2 expression in a mouse
XX demonstrated the presence of two mRNA transcripts (4.8 and 4.4 kb).
XX The levels of these transcripts alter with respect to one another in
XX different tissues. The kidney, spleen and lung appear to express
XX predominantly the larger form, whereas ovary, placenta, skeletal muscle
XX and all murine cell lines analysed express both forms at equal levels.
XX The difference in sizes may be due to differential polyadenylation
XX sites. Both JAK2 and JAK1 are examples of a new subfamily or class
XX of protein tyrosine kinase. These can be used in the phosphorylation
XX of proteins, incorporation of labels and in the design of analogues,
XX antagonists and agonists of JAK's.
XX
XX Sequence 3473 BP; 1128 A; 677 C; 781 G; 887 T; 0 other;
XX
XX Alignment Scores:
XX Pred. No.: 0.0691 Length: 3473
XX Score: 64.00 Matches: 11
XX Percent Similarity: 87.50% Conservative: 3
XX Best Local Similarity: 68.75% Mismatches: 2
XX Query Match: 71.11% Indels: 0
XX DB: Gaps: 0
XX
XX US-09-397-967-15 (1-17) x AAQ25307 (1-3473)
XX
XX QY 2 LysleuleuProLeuAspLysAspTyrTyryValValArgLupProGly 17
XX |||||:||||| |||||:||||| |||||:|||||
XX Db 2566 AAGGCTTGCCACAGACAAAGAAATACATATAAGTAAAGAGCCAGG 2613
XX
XX RESULT 7
XX AAD22680
XX ID AAD22680 standard; cDNA; 3495 BP.
XX
XX AC AAD22680;
XX
XX DT 26-FEB-2002 (first entry)
XX
XX DE Murine JAK2 tyrosine kinase cDNA.
XX
XX DE Murine JAK2 tyrosine kinase cDNA.
XX
XX KW Murine; Src homology 2-Bbeta; SH2-Bbeta; neuroprotective; gene therapy;
XX cell differentiation; nerve regeneration; angiogenesis; embryogenesis;
XX cytosolic; antisense therapy; drug screening; cellular expression;
XX immunological disease; neurological disease; apoptosis; diabetes; cancer;
XX arthritis; JAK2 tyrosine kinase; ss.
XX
XX OS Mus sp.
XX
XX Key Location/Qualifiers
XX FT CDS 1..2982
XX FT /*tag= a
XX FT /product= "Murine JAK2 tyrosine kinase protein"
XX
XX US6312941-B1.
XX
XX PD 06-NOV-2001.
XX
XX XX 26-NOV-1997; 97US-0980080.
XX
XX PF 26-NOV-1997; 97US-0980080.
XX
XX PR 26-NOV-1997; 97US-0980080.
XX
XX (UNMT ) UNIV MICHIGAN.
XX

```


PT Carter-Su C, Rul L, Karow DS;
 XX WPI: 2002-024907/03.
 DR P-PSDB: AA085412.
 XX
 PT New isolated nucleic acid molecule encoding full length Src homology 2
 PT (SH2)-beta protein, useful for modulating cellular expression of
 PT active SH2-beta protein -
 XX
 PS Example 1: Fig 3a: 71pp: English.
 CC The invention relates to human Src homology 2 (SH2)-beta protein and its
 CC DNA molecule. SH2-beta is useful in drug screening assays designed to
 CC identify drugs that interfere with the specific binding of JAK2 tyrosine
 CC kinase which is a member of Janus family of tyrosine kinase-interacting
 CC signalling molecule. SH2-beta protein is useful for modulating,
 CC preferably reducing cellular expression or intracellular concentration or
 CC availability of active SH2-beta. SH2-beta is useful for treating a
 CC disease associated with undesirable cell growth, differentiation, and
 CC growth factor/cytokine, preferably interleukin, more preferably growth
 CC hormone, platelet derived growth factor, nerve growth factor, epidermal
 CC growth factor responsiveness, and for treating diseases associated with
 CC cell movement. SH2-beta molecule has important implications in cancer
 CC metastasis, nerve regeneration, angiogenesis and embryogenesis and is
 CC useful for preventing apoptosis, and treating diabetes, cancer,
 CC arthritis, immunological diseases, neurological diseases, etc.
 CC The invention also relates to compositions and methods for identifying
 CC cytokine, hormone and growth factor signalling pathway agonists and
 CC antagonists. Human SH2-beta DNA is useful in detecting the presence of
 CC SH2-beta genes in gene transcripts, for detecting or amplifying DNA's
 CC with substantial sequence similarity with SH2-beta homologues and
 CC structural analogues and for gene therapy applications. The present
 CC sequence is murine JAK2 tyrosine kinase cDNA related to the invention.
 XX
 SQ Sequence 3495 BP: 1121 A: 688 C: 794 G: 892 T: 0 other:
 Alignment Scores:
 Pred. No.: 0.0696 Length: 3495
 Score: 64.00 Matches: 11
 Percent Similarity: 87.50% Conservative: 3
 Best Local Similarity: 68.75% Mismatches: 2
 Query Match: 71.11% Indels: 0
 DB: 24 Gaps: 0
 US-09-397-967-15 (1-17) x AAD22680 (1-3495)
 QY 2 LysLeuLeuProLeuAspLysAspTyrTyrValValArgGluProGly 17
 DB 2587 AAGCTCTTGCCGACGACAAAGAACTACTCAAGTAAGGACCGAGG 2634
 RESULT 8
 AA085412 standard: cDNA: 3629 BP.
 ID AA085412
 XX
 AC AA085412:
 XX
 DT 05-OCT-1995 (first entry)
 XX
 DE Murine JAK2 kinase coding sequence and flanking regions.
 XX
 KW JAK family; protein tyrosine kinase; cytokine receptor; mouse;
 KW phosphorylation; signal transduction; activation; ss.
 XX
 OS Mus musculus.
 XX
 FH key Location/Qualifiers
 FT CDS 94..3483
 FT /*tag= a
 FT /product= Jak2_kinase
 FT 554..556
 FT /*tag= b
 FT /note= "Published partial sequence of Jak2 cDNA
 FT (Haipur et al., Oncogene 7:1347-1353(1992))"

FT conflict
 FT 1089..3579
 FT /*tag= c
 FT /note= "Published partial sequence of Jak2 cDNA
 FT (Haipur et al., Oncogene 7:1347-1353(1992))"
 FT differs from AA085412 in having the
 FT nucleotides shown in brackets at the following
 FT positions: 1089(T), 1103(C), 1114(G), 1119(G),
 FT 1122(C), 1128(C), 1146(C), 1148(T), 1194(G),
 FT 1140(G), 1143(G), 1146(C), 1188(T), 1272(T),
 FT 1230(A), 1245(T), 1260(T), 1266(C), 1272(T),
 FT 1275(A), 1293(T), 1305(T), 1323(C), 1341(A),
 FT 1344(A), 1359(G), 1365(A), 1368(T), 1374(T),
 FT 1401(C), 1413(C), 1431(T), 1458(A), 1476(G),
 FT 1488(T), 1511..1512(GT), 1578(C), 1590(T),
 FT 1593(T), 1602(G), 1623(T), 1642(G), 1657(C),
 FT 1728(G), 1743(C), 1755(C), 1770(A), 1809(G),
 FT 1816(G), 1821(C), 1857(A), 1878(T), 1935(A),
 FT 1938(A), 1963(T), 1974(G), 2025(T), 2055(G),
 FT 2079(C), 2082(C), 2085(C), 2253(A), 2259(G),
 FT 2283(A), 2285(C), 2433(G), 3453(C), 3579(C)"
 FT
 FT conflict
 FT 2226
 FT /*tag= d
 FT /note= "Location of a 7 amino acid insert detected
 FT in previous studies, but not in the present
 FT study"
 FT
 FT conflict
 FT 3595..3619
 FT /*tag= e
 FT /note= "Three extra nucleotides (all A's) were
 FT present in previous studies at positions
 FT 3595, 3598 and 3619 in the 3'-UTR"
 FT
 FT W09503701-A.
 XX
 PD 09-FEB-1995.
 XX
 PF 29-JUL-1994: 94WO-US08676.
 XX
 PR 29-JUL-1993: 93US-0097997.
 XX
 PA (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.
 XX
 XX Ihle JN, Quelle FW, Silvennoinen O, Witthuhn BA:
 PI WPI: 1995-081950/11.
 DR P-PSDB: AAR70830.
 XX
 PT Inhibiting a cellular response to a cytokine by inhibiting Jak
 PT kinase - to treat diseases caused by excessive response to
 PT cytokine, e.g. erythrocytosis and other cellular proliferative
 PT diseases
 XX
 PS Claim 29; Fig 1; 167pp: English.
 XX
 CC Inhibiting the activity of a Jak kinase (pref. Jak1, Jak2, Jak3 or
 CC Tyk2) in a eukaryotic cell is claimed as a method of inhibiting the
 CC biological response of that cell to a cytokine (not IL-3 or
 CC erythropoietin). The present sequence (murine JAK2 kinase) encodes a
 CC 1129 amino acid protein which includes an epitopic sequence and
 CC amino acid positions 758-776. Antibodies which selectively bind the
 CC epitope are able to bind Jak2 without interfering with the activity
 CC of the kinase. Such antibodies are claimed and are useful for
 CC detecting and extracting Jak2. The 71 nucleotide differences noted
 CC between the present sequence and the published partial sequence
 CC result in 9 amino acid changes.
 XX
 SQ Sequence 3629 BP: 1142 A: 733 C: 842 G: 912 T: 0 other:
 Alignment Scores:
 Pred. No.: 0.0728 Length: 3629
 Score: 64.00 Matches: 11
 Percent Similarity: 87.50% Conservative: 3
 Best Local Similarity: 68.75% Mismatches: 2


```

FT conflict replace (1275, Q)
FT /*tag- x
FT replace (1293, T)
FT /*tag- y
FT replace (1305, T)
FT /*tag- z
FT replace (1323, C)
FT /*tag- aa
FT replace (1341, A)
FT /*tag- ab
FT replace (1344, A)
FT /*tag- ac
FT replace (1359, G)
FT /*tag- ad
FT replace (1365, Q)
FT /*tag- ae
FT replace (1368, T)
FT /*tag- af
FT replace (1374, T)
FT /*tag- ag
FT replace (1401, C)
FT /*tag- ah
FT replace (1413, C)
FT /*tag- ai
FT replace (1431, T)
FT /*tag- aj
FT replace (1453, Q)
FT /*tag- ak
FT replace (1476, G)
FT /*tag- al
FT replace (1488, T)
FT /*tag- am
FT replace (1511..1512, GT)
FT /*tag- an
FT replace (1578, C)
FT /*tag- ao
FT replace (1590, T)
FT /*tag- ap
FT replace (1593, T)
FT /*tag- aq
FT replace (1602, G)
FT /*tag- ar
FT replace (1623, T)
FT /*tag- as
FT replace (1642, G)
FT /*tag- at
FT replace (1657, C)
FT /*tag- au
FT replace (1728, C)
FT /*tag- av
FT replace (1743, C)
FT /*tag- aw
FT replace (1755, C)
FT /*tag- ax
FT replace (1770, A)
FT /*tag- ay
FT replace (1809, G)
FT /*tag- az
FT replace (1816, G)
FT /*tag- ba
FT replace (1821, C)
FT /*tag- bb
FT replace (1857, A)
FT /*tag- bc
FT replace (1878, T)
FT /*tag- bd
FT replace (1935, A)
FT /*tag- be
FT replace (1938, A)
FT /*tag- bf
FT replace (1963, T)
FT /*tag- bg
FT replace (1974, G)
FT conflict

```

```

FT conflict /*tag- bh
FT replace (2025, T)
FT /*tag- bi
FT replace (2055, G)
FT /*tag- bj
FT replace (2079, C)
FT /*tag- bk
FT replace (2082, C)
FT /*tag- bl
FT misc_feature
FT 2226
FT /*tag- bm
FT /note= "Nucleotides corresponding to 7 amino acid insert"
FT replace (2253, A)
FT /*tag- bn
FT replace (2259, G)
FT /*tag- bo
FT replace (2283, A)
FT /*tag- bp
FT replace (2285, C)
FT /*tag- bq
FT replace (2433, G)
FT /*tag- br
FT replace (3453, C)
FT /*tag- bs
FT 3484..3629
FT /*tag- bt
FT replace (3579, C)
FT /*tag- bu
FT replace (3595, A)
FT /*tag- bv
FT replace (3599, A)
FT /*tag- bw
FT replace (3620, A)
FT /*tag- bx

```

```

US6210654-B1.
03-APR-2001.
08-OCT-1997; 97US-0946994.
18-JUN-1996; 96US-0665574.
29-JUL-1993; 93US-0097997.
(SJUD-) ST JUDE CHILDREN'S HOSPITAL.

```

```

Ihle J, Witthuhn BA, Quelle FW, Silvennoinen O;
WPI: 2001-265367/27.
P-PSDB; AAE00352.

```

Modulating a biological response mediated by Jak kinase 2 activation to a cytokine, useful for treating excessive proliferation of eukaryotic cells, comprises inhibiting or enhancing tyrosine kinase activity of Jak kinase in the cell.

Example 1: Fig 1; 100pp; English.

The present sequence is a cDNA encoding murine (Janus kinase 2) Jak2 tyrosine kinase. Jak2 sequence has a 600 amino acid long N-terminus that lacks obvious SH2 (Src homology 2) and SH3 domains. Following this is a kinase related domain (domain 2) and a carboxyl kinase domain (domain 1). Jak kinases mediate cytokine activity through their tyrosine

Alignment Scores:

```

Pred. No.: 0.0728 Length: 3629
Score: 64.00 Matches: 11
Percent Similarity: 87.50% Conservative: 3
Best Local Similarity: 68.75% Mismatches: 2
Query Match: 71.11% Indels: 0
DB: 22 Gaps: 0

```

US-09-397-967-15 (1-17) x AAD03607 (1-3629)

OY 2 LysLeuLeuProLeuAspTyrTyrValValArgLupProGly 17
 |||:::||||| |||:::||||| |||:::|||||
 DB 3088 AAGGCTTGCCGACAGCAAGAACTACTACAAAGTAAAGAGCCAGG 3135

RESULT 11
 AAX80971
 ID AAX80971 standard; cDNA: 4482 BP.
 XX AAX80971;
 AC
 XX 03-SEP-1999 (first entry)
 DT
 XX Human JAK2 kinase encoding cDNA.
 DE
 XX JAK2 kinase; arteriosclerosis; asthma; bronchitis; emphysema; psoriasis;
 KW inflammatory bowel disease; inflammation; osteoarthritis; oncogenesis;
 KW rheumatoid arthritis; septic shock; systemic lupus erythematosus;
 KW leukaemia; human; ss.
 XX
 OS Homo sapiens.
 XX
 PN US5914393-A.
 XX
 PD 22-JUN-1999.
 XX
 PF 05-DEC-1995; 95US-0567508.
 XX
 PR 05-DEC-1995; 95US-0567508.
 XX
 PA (INCY-) INCYTE PHARM INC.
 XX
 PI Coleman R, Stuart SG;
 XX
 DR WPI: 1999-384188/32.
 DR P-PSDB; AAY21698.
 DR
 XX Polypeptides and amino acids useful for modulating human jak2 kinase
 PT activity
 PT
 PS Claim 2; Columns 23-28; 37pp; English.
 XX
 CC This cDNA encodes a human JAK2 kinase polypeptide. Host cells transformed
 CC with recombinant JAK2 kinase nucleic acid are used for the recombinant
 CC production of the protein. Purified JAK2 may be used to produce
 CC antibodies or identify antagonists or inhibitors of JAK2. JAK2, anti-JAK2
 CC antibodies and JAK2 antagonists or inhibitors may be used to treat,
 CC prevent or diagnose conditions associated with altered or uncontrolled
 CC JAK2 expression, e.g. arteriosclerosis, asthma, bronchitis, emphysema,
 CC inflammatory bowel disease, inflammation, leukemia, oncogenesis,
 CC osteoarthritis, psoriasis, rheumatoid arthritis, septic shock and
 CC systemic lupus erythematosus.
 CC
 CC
 SQ Sequence 4482 BP; 1477 A; 795 C; 953 G; 1249 T; 8 other;
 XX

Alignment Scores:
 Pred. No.: 0.0932 Length: 4482
 Score: 64.00 Matches: 11
 Percent Similarity: 87.50% Conservative: 3
 Best Local Similarity: 68.75% Mismatches: 2
 Query Match: 71.11% Indels: 0
 DB: 20 Gaps: 0

US-09-397-967-15 (1-17) x AAX80971 (1-4482)

OY 2 LysLeuLeuProLeuAspTyrTyrValValArgLupProGly 17
 |||:::||||| |||:::||||| |||:::|||||
 DB 3334 AAGGCTTGCCGACAGCAAGAACTACTACAAAGTAAAGAGACCTGCT 3381

RESULT 12
 AAZ58947
 ID AAZ58947 standard; cDNA: 4482 BP.
 XX

AC AAZ58947;
 XX
 DT 03-MAY-2000 (first entry)
 XX
 DE Human JAK2 kinase (HJAK2) encoding cDNA.
 XX
 KW Janus family nonreceptor protein-tyrosine kinase-2; JAK2; HJAK2; human;
 KW signal transduction; arteriosclerosis; asthma; bronchitis; emphysema;
 KW inflammatory bowel disease; leukemia; oncogenesis; osteoarthritis;
 KW psoriasis; rheumatoid arthritis; systemic lupus erythematosus;
 KW cytostatic; osteopathic; dermatological; antibacterial; septic shock;
 KW immunosuppressive; ss.
 XX
 OS Homo sapiens.
 XX
 PN US601966-A.
 XX
 PD 01-FEB-2000.
 XX
 PF 19-NOV-1998; 98US-0196480.
 XX
 PR 05-DEC-1995; 95US-0567508.
 XX
 PA (INCY-) INCYTE PHARM INC.
 XX
 PI Stuart SG, Coleman R;
 XX
 DR WPI: 2000-146859/13.
 DR P-PSDB; AAY77552.
 DR
 XX Human Janus family nonreceptor protein-tyrosine kinases useful as
 PT diagnostic reagents and for preventing, diagnosing and treating
 PT diseases such as arteriosclerosis, asthma and leukemia -
 XX
 PS Example 5; Fig 1A-F; 33pp; English.
 XX
 CC This cDNA encodes a human Janus family nonreceptor protein-tyrosine
 CC kinase-2 (JAK2) polypeptide (HJAK2). The JAK2 polypeptides may be used
 CC as diagnostic reagents as they react with a range of target proteins
 CC including growth hormone, prolactin, erythropoietin and cytokine
 CC receptors. They may also be used for the production of antibodies
 CC specific for JAK2, which may be used to inhibit its activity and prevent
 CC or treat disorders associated with over expression of JAK2. Conversely,
 CC the JAK2 polypeptide may be administered to supplement the patients own
 CC production and counter mutations that may lead to the expression of an
 CC inactive enzyme. The protein may also be used to screen candidate
 CC reagents for modulators of JAK2 function. The antagonists and antibodies
 CC bind to the JAK2 protein and prevent the transfer of high energy
 CC phosphate molecules, therefore blocking signal transduction. Disorders
 CC that may be treated by administration of JAK2 polypeptides, anti-JAK2
 CC antibodies and the agonists and/or antagonists, include arteriosclerosis,
 CC asthma, bronchitis, emphysema, inflammatory bowel disease, leukemia,
 CC oncogenesis, osteoarthritis, psoriasis, rheumatoid arthritis, septic
 CC shock and systemic lupus erythematosus.
 CC
 CC
 SQ Sequence 4482 BP; 1477 A; 795 C; 953 G; 1249 T; 8 other;
 XX

Alignment Scores:
 Pred. No.: 0.0932 Length: 4482
 Score: 64.00 Matches: 11
 Percent Similarity: 87.50% Conservative: 3
 Best Local Similarity: 68.75% Mismatches: 2
 Query Match: 71.11% Indels: 0
 DB: 21 Gaps: 0

US-09-397-967-15 (1-17) x AAZ58947 (1-4482)

OY 2 LysLeuLeuProLeuAspTyrTyrValValArgLupProGly 17
 |||:::||||| |||:::||||| |||:::|||||
 DB 3334 AAGGCTTGCCGACAGCAAGAACTACTACAAAGTAAAGAGACCTGCT 3381

RESULT 13
 AAD24311

-09-397-967-15 (1-17) x AAD24311 (1-5117)

3BL25944
D ABL25944 standard; DNA; 2743 BP

AC ABL25944;

26-MAR-2002 (first entry)

XX DE	Drosophila melanogaster genomic polynucleotide SEQ ID NO 29305.
	1000

XX
Drosophila; developmental biology; cell signalling; insecticide resistance

pharmaceutical; gene; ds.

Drosophila melanogaster.

AA
PN WO200171042-A2.

XX 27-SEP-2001.
PD

XX
PF 23-MAR-2001; 2001WO-US09231.

XX
PR 23-MAR-2000; 2000US-191637P.

11-JUL-2000; 2000US-0614150.

AA (PEKE) PE CORP NY.
PA
NY

XX
PI Venter JC, Adams M, Li PWD, Myers EW;

XX
DR WPI; 2001-656860/75.

XX . New isolated nucleic acid detection reagent for detecting 1000 or more
PT PT all signalling and cell-cell

genes from *Drosophila* and for elucidating cell signalling and cell-cell interactions -

Claim 1: SEO ID NO 29305: 21pp + Sequence Listing; English.

XX relates to an isolated nucleic acid detection

CC The invention relates to an apparatus and a method for
CC capable of detecting 1000 or more genes from *Drosophila*. The invention is
CC directed to the field of biology and in elucidating cell signalling and

CC useful in developmental biology and in characterizing cell-cell interactions in higher eukaryotes for the development of cell-cell interactions in higher eukaryotes for the development of cell-cell interactions in higher eukaryotes for the development of

CC insecticides, therapeutics and pharmaceutical drugs. and expressed DNA
CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA
CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA

sequences (ABL1840-ABL16175) and the encoded proteins (ABR57737-ABR72072).

The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from

CC specification; see www.wipo.int/pub/published_pct_sequences.
CC at ftp.wipo.int/pub/published_pct_sequences.

XX
SQ Sequence 2743 BP; 583 A; 768 C; 739 G; 653 T; 0 other;

Alignment Scores: 2743

Pred. No.:	12.3	Length:	27.5
Score:	51.50	Matches:	11
		Matches:	3

Percent Similarity:	77.78%	Conservative:	3
Percent Local Similarity:	61.11%	Mismatches:	1

Best Local Characteristic:	57.228	Indels:	3
Query Match:	23	Gaps:	1

DB: 20 APR 1964 (1-2743)

US-09-397-967-15 (1-11) X ABL20344 (1-11)

[illegible]

Db 2573 CTTCTCCACTCGTACAGAGTATTACGTTATCCGTGAGTATCTGCTGCTCCGC 202

Search completed: April 28, 2003, 14:00:39
Job time : 13.0636 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 28, 2003, 13:46:54 ; Search time 53.2545 Seconds

(without alignments)
5169.955 Million cell updates/sec

Title: US-09-397-967-15

Perfect score: 90

Sequence: 1 AKLLPLDKDYVVRPG 17

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 segs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters: -DEV-xlp
-MODE=frame+pn.model -DEV-xlp
-O/cgn2.1/USP10.spool/US09397967/runat.26042003.182313.10728/app.query.fasta.1.1486
-DB=EST -OPWT=fastp -SUPFIX=first -MINMATCH=0.1 -LOOPT=0 -LOOPT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blonsum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR=SCORE=pc -THR=MAX=100 -THR=MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTEWT=pc -NOM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09397967 -CGN=1.1.3008 -runat.26042003.182313.10728 -MCPU=6 -ICPU=3
-NO_XLPXY -NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST: *
1: em_estba: *
2: em_estlum: *
3: em_estlum: *
4: em_estlum: *
5: em_estlum: *
6: em_estlum: *
7: em_estlum: *
8: em_estlum: *
9: gb_est1: *
10: gb_est1: *
11: gb_est1: *
12: gb_est1: *
13: gb_est1: *
14: gb_est1: *
15: em_estlum: *
16: em_estlum: *
17: gb_est1: *
18: em_gss_hum: *
19: em_gss_hum: *
20: em_gss_hum: *
21: em_gss_hum: *
22: em_gss_hum: *
23: em_gss_hum: *
24: em_gss_hum: *
25: em_gss_hum: *
26: em_gss_hum: *
27: em_gss_hum: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	90	100.0	462	10	AM408832
2	90	100.0	544	12	BF705690
3	90	100.0	562	12	BF193178
4	86	95.6	417	13	BF1540725
5	86	95.6	481	17	AO192794
6	86	95.6	543	12	BE753801
7	83	92.2	453	9	AA023709
8	83	92.2	478	9	AA023670
9	83	92.2	512	9	AA010093
10	83	92.2	593	9	AA081653
11	83	92.2	698	14	BM935182
12	83	92.2	720	12	BM873355
13	83	92.2	907	12	BM870513
14	80	88.9	592	9	A1981562
15	80	88.9	769	9	AJ442008
16	74	82.2	590	14	W48204
17	69	76.7	989	17	CNS042XQ
18	69	76.7	1001	17	CNS038MW
19	67	74.4	449	10	BB711016
20	66	73.3	867	9	AA755769
21	64	71.1	398	9	AA53345
22	64	71.1	532	12	BM364706
23	64	71.1	562	10	AM231037
24	64	71.1	753	12	BM870416
25	64	71.1	767	13	BM734179
26	64	71.1	920	14	BM898824
27	64	71.1	945	14	BM732700
28	64	71.1	945	14	BM732700
29	58	64.4	258	12	BF367544
30	57	63.3	612	9	AU091091
31	57	63.3	694	10	AA521703
32	51.5	57.2	687	9	AA699477
33	51.5	57.2	709	13	AA539009
34	51	56.7	141	9	B1588497
35	51	56.7	308	9	AA014812
36	51	56.7	434	17	AA174246
37	51	56.7	766	13	AA0177309
38	49	54.4	379	10	AA540081
39	49	54.4	437	17	AA009280
40	49	54.4	551	12	BF194671
41	49	54.4	569	17	AA014068
42	49	54.4	651	10	BB620498
43	49	54.4	789	10	AA154977
44	48.5	53.9	851	17	CNS03HSP
45	48	53.3	319	9	AA358653

ALIGNMENTS

RESULT 1
AM408832
LOCUS: AM408832
DEFINITION: UT-HF-BMO-idx-h-03-0-UT-rl NIH-MGC-38 Homo sapiens cDNA clone
IMAGE: 3063389 5', mRNA sequence.
ACCESSION: AM408832
VERSION: AM408832.1 GI:6927889
KEYWORDS: human.
SOURCE: Homo sapiens
ORGANISM: Homo sapiens
REFERENCE: Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS: Mammalia; Euthera; Primates; Catarrhini; Homiidae; Homo.
TITLE: NIH-MGC http://mgc.ncbi.nlm.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL
COMMENT

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: rgs@pshs-remail.nih.gov
Eco RI site shown at the beginning of the sequence.
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
CDNA Library Preparation: M.B. Soares Lab
CDNA Library Arrayed by: M.B. Soares Lab
DNA Sequencing by: M.B. Soares Lab
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E.E. Consortium/ILM at:
www.bio.liml.gov/bdrp/image/image.html
Seq primer: M13 Forward.

FEATURES
source

Location/Qualifiers
1..462

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="3063389"
/clone_lib="NIH_MGC_38"
/tissue_type="lymph"
/cell_type="germinal center B cells"
/cell_line="MGC85"
/lab_host="DH10B (LTR)"
/note="Vector: pTRT3-Pac; Site_1: NotI; Site_2: Eco RI;
Constructed from size fractionated cytoplasmic mRNA
(2.5-3.5kb). Directionally cloned. Cells provided by Louis
M. Staudt, Ph.D. Library preparation by Maria de Fatima
Bonaldi, Ph.D. and M. Bento Soares, Ph.D."

BASE COUNT
ORIGIN

86 a 156 c 131 g 89 t

Alignment Scores:

Pred. No.: 2.24e-06 Length: 462
Score: 90.00 Matches: 17
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-09-397-967-15 (1-17) x AM408832 (1-462)

Oy 1 AlalysleuProleuaspyspyrryValaValaArgluProgly 17

Db 349 GCTAGCTGCTCCGCTGACAAAGACTACTAGTGTGCGGACCGACGCG 399

RESULT 2

LOCUS BF705690 544 bp mRNA linear EST 29-DEC-2000
DEFINITION 243606 MARC 2P1G Sus scrofa cDNA 5', mRNA sequence.
ACCESSION BF705690
VERSION BF705690.1 GI:11997351
KEYWORDS EST.
SOURCE pig.
ORGANISM Sus scrofa

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.

REFERENCE 1 (bases 1 to 544)
Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E.,
Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Laegreid,W.W.
and Keeler,J.W.

Design and use of two pooled tissue normalized cDNA libraries for
EST discovery in swine

Unpublished (2000)
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390

JOURNAL
COMMENT

Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -minmatch 12 options.
PCR primers
FORWARD: AGGAACAGCTATGACCAT

BACKWARD: GTTTTCCAGTCACGACG
Plate: 70 row: D column: 5
Seq primer: ATTAGGAGACACTATAG.

FEATURES
source

Location/Qualifiers
1..544

/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone_lib="MARC 2P1G"
/tissue_type="pooled"
/lab_host="DH10B"
/note="Vector: pCMV SPORT6; Site_1: NotI; Site_2: SalI;
Library made from pooled tissue from testis, ovary,
endometrium, hypothalamus, pituitary, and placenta."

BASE COUNT
ORIGIN

96 a 181 c 156 g 111 t

Alignment Scores:

Pred. No.: 2.82e-06 Length: 544
Score: 90.00 Matches: 17
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

US-09-397-967-15 (1-17) x BF705690 (1-544)

Oy 1 AlalysleuProleuaspyspyrryValaValaArgluProgly 17

Db 154 GCGAGCTGCTCCGCTGACAAAGACTACTATGTGTGCGGACCGCGCG 204

RESULT 3

LOCUS BF193178 562 bp mRNA linear EST 02-NOV-2000
DEFINITION 244670 MARC 2P1G Sus scrofa cDNA 5', mRNA sequence.
ACCESSION BF193178
VERSION BF193178.1 GI:11076547
KEYWORDS EST.
SOURCE pig.
ORGANISM Sus scrofa

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.

REFERENCE 1 (bases 1 to 562)
Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Laegreid,W.W.
and Keeler,J.W.

Design and use of two pooled tissue normalized cDNA libraries for
EST discovery in swine

Unpublished (2000)
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390

JOURNAL
COMMENT

Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -minmatch 12 options.
PCR primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTTCCAGTCACGACG
Plate: 72 row: F column: 7
Seq primer: ATTAGGAGACACTATAG.

FEATURES
source

Location/Qualifiers
1..562

/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone_lib="MARC 2P1G"
/tissue_type="pooled"
/lab_host="DH10B"
/note="Vector: pCMV SPORT6; Site_1: NotI; Site_2: SalI;
Library made from pooled tissue from testis, ovary,
endometrium, hypothalamus, pituitary, and placenta."

BASE COUNT

97 a 188 c 167 g 110 t

ORIGIN

Alignment Scores:

Pred. No.: 2,95e-06 Length: 562

Score: 90.00 Matches: 17

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 12 Gaps: 0

US-09-397-967-15 (1-17) x BF193178 (1-562)

Oy 1 AlalysleuProleuAsplysAspyrTyValValArgJupProgly 17

Db 165 GCGAAGCTGCTGCCCTGACAAAGACTACTATGTGTGCGGAGCCGCCGC 215

RESULT 4

LOCUS B1540725 417 bp mRNA linear EST 30-AUG-2001

DEFINITION 4541025 MARC 1BOV Bos taurus cDNA 5', mRNA sequence.

ACCESSION B1540725

VERSION B1540725.1 GI:15381837

KEYWORDS EST.

SOURCE

ORGANISM Bos taurus

COV.

REFERENCE

AUTHORS 1 (bases 1 to 417)

Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T., Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-McKown,C.G., Pettea,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,J. and Keefe,J.W.

TITLE Sequence evaluation of four pooled-tissue normalized bovine cDNA libraries and construction of a gene index for cattle

JOURNAL Genome Res. 11 (4), 626-630 (2001)

MEDLINE 21180013

COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov

Single pass sequencing. Bases called and alt-trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 18 and -mismatch 12 options.

PCR primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTCCAGTCACGACG
Plate: 118 row: P column: 4
Seq primer: ATTAGGTGACACTTAG.

FEATURES

source

1..417

/organism="Bos taurus"

/db_xref="taxon:9913"

/clone_lib="MARC 1BOV"

/tissue_type="pooled"

/lab_host="DH10B"

/note="Vector: pCMV SPORT6; Site_1: NotI; Site_2: SalI; library made from pooled tissue from lymph node, ovary, fat, hypothalamus, and pituitary."

BASE COUNT 74 a 138 c 121 g 83 t 1 others

ORIGIN

Alignment Scores:

Pred. No.: 1.04e-05 Length: 417

Score: 86.00 Matches: 16

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 94.12% Mismatches: 1

Query Match: 95.56% Indels: 0

DB: 13 Gaps: 0

US-09-397-967-15 (1-17) x B1540725 (1-417)

Oy 1 AlalysleuProleuAsplysAspyrTyValValArgJupProgly 17

Db 85 GCGAAGCTGCTGCCCTGACAAAGACTACTATGTGTGCGGAGCCGCCGC 135

RESULT 5

LOCUS A0192794/c 481 bp DNA linear GSS 04-NOV-1998

DEFINITION HS-2251_B2_F07_MR CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate-2251 Col-14 Row-L, DNA sequence.

ACCESSION A0192794

VERSION A0192794.1 GI:3591416

KEYWORDS GSS.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE

AUTHORS 1 (bases 1 to 481)

Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.

TITLE Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)

MEDLINE 99380589

COMMENT Contact: Mahairas Gc, Wallace Jc, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu

Sequence Tagged Connector
Plate: 2251 row: L column: 14
Class: BAC ends
High quality sequence stop: 481.

FEATURES

source

1..481

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="Plate-2251 Col-14 Row-L"

/sex="male"

/note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in E-Coli DH10B"

BASE COUNT 101 a 115 c 192 g 72 t 1 others

ORIGIN

Alignment Scores:

Pred. No.: 1.27e-05 Length: 481

Score: 86.00 Matches: 16

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 95.56% Indels: 0

DB: 17 Gaps: 0

US-09-397-967-15 (1-17) x A0192794 (1-481)

Oy 2 LysleuProleuAsplysAspyrTyValValArgJupProgly 17

Db 360 AACCTGCTGCCCTGACAAAGACTACTATGTGTGCGGAGCCGCCGC 313

RESULT 6

LOCUS BE753801 543 bp mRNA linear EST 25-APR-2001

DEFINITION 206963 MARC 2BOV Bos taurus cDNA 5', mRNA sequence.

ACCESSION BE753801

VERSION BE753801.1 GI:10167793

KEYWORDS EST.

SOURCE cow.

ORGANISM Bos taurus

REFERENCE

AUTHORS 1 (bases 1 to 543)

Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T., Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-McKown,C.G., Pettea,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,J. and Keefe,J.W.

TITLE Sequence evaluation of four pooled-tissue normalized bovine cDNA libraries and construction of a gene index for cattle

JOURNAL Genome Res. 11 (4), 626-630 (2001)

MEDLINE 21180013

COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov

Single pass sequencing. Bases called and alt-trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 18 and -mismatch 12 options.

PCR primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTCCAGTCACGACG
Plate: 118 row: P column: 4
Seq primer: ATTAGGTGACACTTAG.

FEATURES

source

1..543

/organism="Bos taurus"

/db_xref="taxon:9913"

/clone_lib="MARC 2BOV"

/tissue_type="pooled"

/lab_host="DH10B"

/note="Vector: pCMV SPORT6; Site_1: NotI; Site_2: SalI; library made from pooled tissue from lymph node, ovary, fat, hypothalamus, and pituitary."

BASE COUNT 74 a 138 c 121 g 83 t 1 others

ORIGIN

Alignment Scores:

Pred. No.: 1.04e-05 Length: 417

Score: 86.00 Matches: 16

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 94.12% Mismatches: 1

Query Match: 95.56% Indels: 0

DB: 13 Gaps: 0

REFERENCE
AUTHORS
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 543)
Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
Casas,E., Wray,D.E., White,J., Cho,U., Fahrenkrug,S.C., Bennett,
G.L., Heaton,M.P., Laegreid,W.M., Rohrer,G.A., Chitko-McKown,C.G.,
Pettee,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,J. and
Keele,J.W.
Sequence evaluation of four pooled-tissue normalized bovine cDNA
libraries and construction of a gene index for cattle
Genome Res. 11 (4), 626-630 (2001)
21180013
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smithemall.marc.usda.gov
Single pass sequencing. Bases called and alt_trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -minmatch 12 options.
PCR primers
FORWARD: AGGAACAGCATGACCAT
BACKWARD: GTTTCACAGTCACGACG
Plate: 49 row: M column: 16
Seq primer: ATTAGCTGACACTATAG.
Location/Qualifiers
1. 543
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone_lid="MARC 2BOY"
/rissue_type="pooled"
/lab_host="DH10B"
/note="Vector: PCMV SPORT6; Site_1: NotI; Site_2: SalI;
library made from pooled tissue from testis, thymus,
semilandonous muscle, longissimus muscle, pancreas,
adrenal, and endometrium."

BASE COUNT
ORIGIN
100 a 182 c 158 g 103 t

Alignment Scores:
Pred. No.: 1.51e-05 Length: 543
Score: 86.00 Matches: 16
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 94.12% Mismatches: 0
Query Match: 95.56% Indels: 0
DB: 12 Gaps: 0

US-09-397-967-15 (1-17) x BE753801 (1-543)

QY 1 AAlaAysleuLeuProLeuAspLysAspTyrTyValValArgGIupProGly 17
|||||
Db 333 GCCAAGCTGCTGCTCGACAAAGAGTACTGCTGCTCCGCGAGCCAGGC 383
|||||

RESULT 7
AA023709 453 bp mRNA linear EST 21-JAN-1997
LOCUS mh77602.f1 Soares mouse placenta 4NDMP13.5 14.5 Mus musculus cDNA
DEFINITION clone IMAGE:456986 5' similar to gp:L32955 Mouse protein tyrosine
kinase (MOUSE);, mRNA sequence.
AA023709
AA023709.1 GI:1487624
EST.
house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 453)
Marr,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Thelning,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.

REFERENCE
AUTHORS

TITLE
JOURNAL
COMMENT
The Washu-HHMI Mouse EST Project
Unpublished (1996)
Contact: Marra M/Mouse EST Project
Washu-HHMI Mouse EST Project
Washington University School of Medicine
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LIND; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:273874
Seq primer: -28M13 rev2 from Amersham
High quality sequence stop: 406.
Location/Qualifiers
1. 453
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone_lid="IMAGE:456986"
/clone_lid="Soares mouse placenta 4NDMP13.5 14.5"
/sex="unknown"
/rissue_type="placenta"
/dev_stage="adult"
/lab_host="DH10B"
/note="Organ: placenta; Vector: pT73D-Pac (Pharmacia)
with a modified polylinker; Site_1: Not I; Site_2: Eco RI;
1st strand cDNA was primed with a Not I - oligo(dt) primer
15'
TGTTACCATCTGTAAGTGGAGCGCCGCGAAATTTTTTTTTTTTTTTTTTTT
T 3'; double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M.Fatima Bonaldo."

BASE COUNT
ORIGIN
79 a 153 c 137 g 84 t

Alignment Scores:
Pred. No.: 4.13e-05 Length: 453
Score: 83.00 Matches: 16
Percent Similarity: 94.12% Conservative: 0
Best Local Similarity: 94.12% Mismatches: 1
Query Match: 92.22% Indels: 0
DB: 9 Gaps: 0

US-09-397-967-15 (1-17) x AA023709 (1-453)

QY 1 AAlaAysleuLeuProLeuAspLysAspTyrTyValValArgGIupProGly 17
|||||
Db 30 GCTAAGCTGCTGCCCTGGGAAAGAGTACTGCTGCTCCGCGAGCTGGC 80
|||||

RESULT 8
AA023670 478 bp mRNA linear EST 21-JAN-1997
LOCUS mh78606.f1 Soares mouse placenta 4NDMP13.5 14.5 Mus musculus cDNA
DEFINITION clone IMAGE:457114 5' similar to gp:L32955 Mouse protein tyrosine
kinase (MOUSE);, mRNA sequence.
AA023670
AA023670.1 GI:1487587
EST.
house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 478)
Marr,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Thelning,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.

REFERENCE
AUTHORS

TITLE
JOURNAL
The Washu-HHMI Mouse EST Project
Unpublished (1996)

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

Email: mouseest@atson.wustl.edu
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:667642

Seq primer: -28m13 rev2 ET from Amersham
High quality sequence stop: 526.

FEATURES

Source

```

/jorganism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:1265090"
/clone_id="Soares_thymus_2NbMT"
/sex="male"
/tissue_type="Thymus"
/dev_stage="4 weeks"
/lab_host="DH10B"
/notes="vector: pT7.2-Pac (Pharmacia) with a modified
polylinker. Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was primed with a Not I oligo(dT) primer 15',
3'TTACCAATGTCAGTGAAGGAGGAGCGCCGCTTTTTTT
TTTTTTT
TTTT. double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT7.3 vector. RNA
provided by Dr. Bertrand Jordan. Library went through two
rounds of normalization, and was constructed by Bento
Soares and M.Fatima Donaldo."
124.†

```

BASE COUNT	106 a	197 c	166 g	124. l
ORIGIN				

ORIGIN

Alignment Scores:

Alignment Score:	6.06e-05	length:	593
Pred. NO.:	83.00	Matches:	16
Score:	94.12%	Conservative:	0
Percent Similarity:	94.12%	Mismatches:	1
Best local Similarity:	92.22%	Indels:	0
Query Match:	9	Gaps:	0
DB:			

US-09-397-967-15 (1-17) x AA881653 (1-593)

QY 1 AAlalysleuleuproleuaspIysasPTyrrTyValValarGluProgly 17
 |||||
Db 39 GCTAAGCTGCTGCCCTGGGAAGAAGACTACTACGTGTCGCCGAGCCTGCC 89

RESULT 11
BM935182
c08 bo
mdna linear
EST 13-MAR-2002

LOCUS	BM935182	698 bp	mRNA	linear	EST 13-MAR-
DEFINITION	UI-M-BH3-ag2-f-08-0-UI.r1 NIH_BMAP_M_S4 Mus musculus cDNA clone				
	UI-M-BH3-ag2-f-08-0-UI 5', mRNA sequence.				

BM935182.1 GI:19394334

SOURCE house mouse.

REFERENCE	
AUTHORS	
TITLE	
1 (bases 1 to 698)	
Bonaldo,M.F., Lennon,G. and Soares,M.B.	
Normalization and subtraction: two approaches to facilitate gene	

JOURNAL
Genome Res. 6 (9), 791-806 (1996)
07044477

COMMENT

0001 EXECUTIVE DIV.: ROOM 701, 1200 N. 17TH AVE.,
20892-9643, USA

Fax: 301 443 9890

CDNA Library preparation: Dr. M. Bento Soares, University of Iowa

CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
Accession Number: U01289

Seq primer: M13 REVERSE.

FEATURES	Location/Qualifiers
source	1. .698

/organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="U1-M-BH3-agf-f-08-0-01"
 /clone_id="NIH_BMAP_M_S4"
 /dev_stage="27-32 days"
 /note="vector: p7YHD-Pac (Pharmacia) with a modified
 polylinker. Site.1: Not I; Site.2: Eco RI; The
 NIH_BMAP_M_S4 library is a subtracted library of a series,
 ultimately derived from a mixture of individually tagged
 normalized libraries from ten regions of the mouse brain
 (cerebellum, brain stems, olfactory bulbs, hypothalamus,
 cortex, amygdala, basal ganglia, pineal gland, striatum,
 hippocampus) after a series of subtractions to reduce the
 representation of cDNAs from which ESTs had already been
 generated. The following serially subtracted libraries
 were generated in this process: NIH_BMAP_M_S4,
 NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, NIH_BMAP_M_S3.1,
 NIH_BMAP_M_S2, NIH_BMAP_M_S1. The subtracted library
 (NIH_BMAP_M_S4) was constructed as follows: PCR-amplified
 cDNA inserts from NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, and
 NIH_BMAP_M_S3.1 clones from which 3 ESTs had been delivered
 was used as a driver in a hybridization with a pool of
 the NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, and NIH_BMAP_M_S3.1
 libraries in the form of single-stranded circles. The
 remaining single-stranded circles (subtracted library)
 was purified by hydroxyapatite column chromatography,
 converted to double-stranded circles and electroporated
 into DH10B bacteria (Life Technologies) to generate the
 NIH_BMAP_M_S4 library. This procedure has been previously
 described (Donald, Lennon and Soares, Genome Research
 6:791-806, 1996)".

BASE COUNT	132 a	211 c	215 g	138 t	2 others
ORIGIN					

ORIGIN

Alignment Scores:

Pred. No.:	7,048	83.00
Score:	85.00	1
Percent Similarity:	94.12%	0
Best local Similarity:	94.12%	1
Query Match:	92.22%	0
DB:	14	0
Gaps:		0

US-09-397-967-15 (1-17) x BM935182 (1-698)

1 AlaLysLeuProLeuAspLysAspTyrTyrValValArgGluProGly 17

Db 525 GCTAGCTGCTGCCCTGGGAAGAGACTACTACGTGGTCCGAGACCTGGC 575

RESULT 12
BG873355
DEC 30-WAY-2001

LOCUS	BG873335	720 bp	mRNA	linear	EST:23-MAR-2004
DEFINITION	602794228F1	NCI_CGAP_SG2	MUS	musculus	CDNA clone IMAGE:4925598 5'

mRNA sequence, BG873355
 ACCESSION

VERSION BG873355.1 GI:14223893
KEYWORDS EST.
SOURCE house mouse

SOURCE
house mouse.
ORGANISM
Mus musculus
Mus musculus
Fukuyama, Metazoa: Chordata: Craniata; Vertebrata; Euteleostomi;

REFERENCE
1. (bases 1 to 720)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

COMMENT

Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM10850 row: g column: 07
High quality sequence stop: 709.

FEATURES

source

1..720
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4925598"
/lab_host="NCI CGAP_SG2"
/note="Organ: salivary gland; Vector: PCMV-SPORT6; Site: 1; NCI; Site_2: Salt; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.3 kb. Constructed by Life Technologies. Note: this is a NCI CGAP Library."
BASE COUNT 123 a 239 c 229 g 129 t
ORIGIN

Alignment Scores:

Pred. No.: 7.99e-05 Length: 720
Score: 83.00 Matches: 16
Percent Similarity: 94.12% Conservative: 0
Best Local Similarity: 94.12% Mismatches: 1
Query Match: 92.22% Indels: 0
DB: 12 Gaps: 0

US-09-397-967-15 (1-17) x BG87355 (1-720)

QY 1 AlalysleuLeuProLeuAspTyTYrValValArgGluProGly 17
|||||
Db 229 GCTAAGCTGCTGCCCTGGGAAGACTACTAGCTGCCGAGCCTGCC 279

RESULT 13

BG870513

LOCUS

602791432F1 NCI CGAP_SG2 Mus musculus cDNA clone IMAGE:4922742 5',
DEFINITION

602791432F1 NCI CGAP_SG2

602791432F1 NCI CGAP_SG2

602791432F1 NCI CGAP_SG2

602791432F1 NCI CGAP_SG2

602791432F1 NCI CGAP_SG2

602791432F1 NCI CGAP_SG2

602791432F1 NCI CGAP_SG2

602791432F1 NCI CGAP_SG2

602791432F1 NCI CGAP_SG2

602791432F1 NCI CGAP_SG2

602791432F1 NCI CGAP_SG2

602791432F1 NCI CGAP_SG2

602791432F1 NCI CGAP_SG2

602791432F1 NCI CGAP_SG2

602791432F1 NCI CGAP_SG2

602791432F1 NCI CGAP_SG2

602791432F1 NCI CGAP_SG2

602791432F1 NCI CGAP_SG2

602791432F1 NCI CGAP_SG2

602791432F1 NCI CGAP_SG2

602791432F1 NCI CGAP_SG2

602791432F1 NCI CGAP_SG2

602791432F1 NCI CGAP_SG2

602791432F1 NCI CGAP_SG2

602791432F1 NCI CGAP_SG2

602791432F1 NCI CGAP_SG2

602791432F1 NCI CGAP_SG2

602791432F1 NCI CGAP_SG2

602791432F1 NCI CGAP_SG2

602791432F1 NCI CGAP_SG2

602791432F1 NCI CGAP_SG2

602791432F1 NCI CGAP_SG2

/clone.lib="NCI CGAP_SG2"

/lab_host="DH10B (T1 phage-resistant)"

/note="Organ: salivary gland; Vector: PCMV-SPORT6; Site: 1; NCI; Site_2: Salt; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.3 kb. Constructed by Life Technologies. Note: this is a NCI CGAP Library."

BASE COUNT 166 a 279 c 275 g 185 t
ORIGIN

Alignment Scores:

Pred. No.: 0.000111 Length: 907
Score: 83.00 Matches: 16
Percent Similarity: 94.12% Conservative: 0
Best Local Similarity: 94.12% Mismatches: 1
Query Match: 92.22% Indels: 0
DB: 12 Gaps: 0

US-09-397-967-15 (1-17) x BG870513 (1-907)

QY 1 AlalysleuLeuProLeuAspTyTYrValValArgGluProGly 17
|||||
Db 366 GCTAAGCTGCTGCCCTGGGAAGACTACTAGCTGCCGAGCCTGCC 416

RESULT 14

AI981562

LOCUS

592 bp mRNA linear EST 07-MAY-2001
pat.pk0060.a5.f chicken activated T cell cDNA Gallus gallus cDNA
clone pat.pk0060.a5.f 5' similar to JAK3 TYROSINE-PROTEIN KINASE,
DEFINITION

592 bp mRNA linear

592 bp mRNA linear

592 bp mRNA linear

592 bp mRNA linear

592 bp mRNA linear

592 bp mRNA linear

592 bp mRNA linear

592 bp mRNA linear

592 bp mRNA linear

592 bp mRNA linear

592 bp mRNA linear

592 bp mRNA linear

592 bp mRNA linear

592 bp mRNA linear

592 bp mRNA linear

592 bp mRNA linear

592 bp mRNA linear

592 bp mRNA linear

592 bp mRNA linear

592 bp mRNA linear

592 bp mRNA linear

592 bp mRNA linear

592 bp mRNA linear

592 bp mRNA linear

592 bp mRNA linear

592 bp mRNA linear

592 bp mRNA linear

592 bp mRNA linear

592 bp mRNA linear

592 bp mRNA linear

592 bp mRNA linear

592 bp mRNA linear

592 bp mRNA linear

592 bp mRNA linear

592 bp mRNA linear

592 bp mRNA linear

592 bp mRNA linear

592 bp mRNA linear

592 bp mRNA linear

592 bp mRNA linear

592 bp mRNA linear

592 bp mRNA linear

592 bp mRNA linear

592 bp mRNA linear

592 bp mRNA linear

592 bp mRNA linear

592 bp mRNA linear

US-09-397-967-15 (1-17) x AI981562 (1-592)

QY 1 AlalysleuLeuProLeuAspTyTYrValValArgGluProGly 17
|||||
Db 366 GCTAAGCTGCTGCCCTGGGAAGACTACTAGCTGCCGAGCCTGCC 416

BASE COUNT 113 a 177 c 184 g 109 t
ORIGIN

FEATURES

source

1..592
/organism="Gallus gallus"
/db_xref="taxon:9031"
/clone="pat.pk0060.a5.f"
/clone.lib="chicken activated T cell cDNA"
/sex="male"
/cell_type="Con A-activated splenic T cell"
/lab_host="E.coli TOP10 F"
/note="Vector: pCDNA3"

Alignment Scores:

Pred. No.: 0.000214 Length: 592
Score: 80.00 Matches: 15
Percent Similarity: 94.12% Conservative: 1
Best Local Similarity: 88.24% Mismatches: 1
Query Match: 88.89% Indels: 0
DB: 9 Gaps: 0

Oy 1 AlalysleuLeuProLeuAspTyrValValArgGluProGly 17
 ||||||||||||||| |||||||||||||||
 Db 101 GCCAAGCTGCTGCCAGACAGAGACTACTACTGCTGCGAGAGCCCGG 151

RESULT 15

AJ442008 769 bp mRNA linear EST 19-APR-2002
 LOCUS AJ442008 dkfz426 Gallus gallus cDNA clone 17j3r1, mRNA sequence.

DEFINITION

AJ442008

ACCESSION

AJ442008.1 GI:20209229

VERSION

EST.

KEYWORDS

chicken.

SOURCE

Gallus gallus

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 Phasianinae; Gallus.

REFERENCE

1 (bases 1 to 769)

AUTHORS

Buerstedde,J.M.

TITLE

Gallus gallus bursal lymphocyte EST

JOURNAL

Unpublished (2002)

COMMENT

Contact: Buerstedde JM
 Cellular Immunology
 Heinrich-Pette-Institute
 Martinistr. 52, 20251 Hamburg, Germany
 Email: URL: <http://genetics.npi.uni-hamburg.de/dt40est.html>.

FEATURES

source

1..769

/organism="Gallus gallus"

/strain="C8"

/db_xref="taxon:9031"

/clone="17j3r1"

/clone_lib="dkfz426"

/tissue_type="Bursa of Fabricius"

/cell_type="bursal lymphocyte"

/dev_stage="2-3 weeks old"

BASE COUNT 150 a 222 c 262 g 134 t 1 others

ORIGIN

Alignment Scores:

Pred. No.: 0.00031 Length: 769

Score: 80.00 Matches: 15

Percent Similarity: 94.12% Conservative: 1

Best Local Similarity: 88.24% Mismatches: 1

Query Match: 88.89% Indels: 0

DB: 9 Gaps: 0

US-09-397-967-15 (1-17) x AJ442008 (1-769)

QY 1 AlalysleuLeuProLeuAspTyrValValArgGluProGly 17

|||||

Db 589 GCCAAGCTGCTGCCAGACAGAGACTACTACTGCTGCGAGAGCCCGG 639

Search completed: April 28, 2003, 16:55:03
 Job time : 57.2545 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 28, 2003, 13:46:54 ; Search time 3442.75 Seconds

(without alignments)
5169.955 Million cell updates/sec

Title: US-09-397-967-16

Perfect score: 5860

Sequence: 1 MAPSEPTPLIPORCSLSS.....RPAPATLSPQDPLMRGPG 1099

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

```

-MODEL=frame_plus_p2n_model -DEV=xlp
-O=/cg92.1/USPRO.spool/US09397967/cunat_26042003_182313_10728/app.query.fasta.1.1486
-DB=EST -QFMT=fastcap -SUFFIX=first -MINMATCH=0.1 -LOOFCLE=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=biolum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09397967 -CGN=1.1.3008 -runat_26042003_182313_10728 -NCPU=6 -ICPU=3
-NO_XLIPY -NO_MMAP -LARGEQUERY -NEG.SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FCGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

```

Database:

```

EST:
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rnd:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1380	23.5	971	13	B1411962
2	1358	23.2	1625	11	BC027234
3	1192	20.3	628	13	B1557690
4	1141	19.5	898	14	BM935182
5	1134	19.4	675	12	BF150250
6	1031.5	17.6	907	12	BC870513
7	1028.5	17.6	720	12	BC873355
8	1016	17.3	621	10	AW742526
9	1012	17.3	731	9	AA920299
10	960	16.4	583	10	BB631706
11	935.5	16.0	871	14	BQ709825
12	930	15.9	559	12	BC800853
13	906	15.5	769	9	AJ442008
14	879	15.0	545	12	BE916148
15	874	14.9	539	12	BE913247
16	835.5	14.3	543	12	BE753801
17	831.5	14.2	562	12	BF193178
18	828	14.1	580	10	BB608771
19	824	14.1	557	13	BI338676
20	823	14.0	544	12	BF705690
21	807	13.8	590	14	W48204
22	782.5	13.4	867	9	AA755769
23	781.5	13.3	554	12	BF191625
24	762	13.0	767	13	BI734179
25	761	13.0	763	9	AJ399251
26	759.5	13.0	564	13	BM088287
27	725.5	12.4	540	10	AW505513
28	723	12.3	925	14	BO683553
29	721	12.3	1097	13	BM455276
30	720.5	12.3	526	12	BF191620
31	709.5	12.1	593	9	AA881653
32	705	12.0	501	10	BE623895
33	702	12.0	436	10	AW141115
34	698	11.9	453	9	AA023709
35	695.5	11.9	462	10	AA408832
36	693	11.9	512	9	AA510093
37	693	11.8	478	9	AA023670
38	680.5	11.6	737	9	AU134500
39	666.5	11.4	491	12	BF441858
40	651.5	11.1	923	9	AU133110
41	644.5	11.0	550	10	BE309453
42	638.5	10.9	562	10	AW231037
43	637.5	10.9	472	10	AW785501
44	633.5	10.8	917	12	BG767031
45	628.5	10.7	689	12	BF614255

ALIGNMENTS

RESULT 1
LOCUS B1411962
DEFINITION 602966375P1 NCI_CGAP_Lu33 Mus musculus cDNA clone IMAGE:5121786 5',
B1411962
ACCESSION B1411962
VERSION B1411962.1 GI:15172885
KEYWORDS EST.
ORGANISM house mouse.
MUS MUSCULUS
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 971)
NIH-MGC http://mgc.nci.nih.gov/
TITLES National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL COMMENT

Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Gilbert Smith, Ph.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
 Plate: LLM11297 row: e column: 19
 High quality sequence start: 25
 High quality sequence stop: 785.
 Location/Qualifiers

FEATURES

1. 971
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="IMAGE:5121786"
 /clone_1lb="NCL_CGAP_Lu33"
 /tissue_type="pooled lung tumors"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: lung; Vector: pTZ19-pac (Pharmacia) with a modified polylinker; Site_1: NotI; Site_2: EcoRI; 1st strand cDNA was prepared from mRNA obtained from pooled lung tumors with a Not I - oligo(dT) primer [5' TGTTCACATCTGACAGTGGAGGCGCCCTGTTTCTTTTCTTTT 3']
 Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pTZ19 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 216 a 274 c 279 g 202 t

ALIGNMENT SCORES:

Score: 1.43e-129 Length: 971
 Percent Similarity: 1380.00 Matches: 287
 Best Local Similarity: 89.30% Conservative: 5
 Query Match: 87.77% Mismatches: 27
 Query Match: 23.55% Indels: 10
 Gaps: 2

US-09-397-967-16 (1-1099) x B1411962 (1-971)

QY 475 ProlysgluysSeranleuValaValaArgglyCysAsnProalaProalaPro 494
 Db 6 CCAAGGTCAGAGTATTTTGTATGATGTGGAGAGGGCTGACCCCGCCCTGCCCT 65
 QY 495 GlyCysSerProSerCysAlaLeuThrcLeuSerPheHisThrIleProThrAsp 514
 Db 66 GGGTGGTCCCGCTCCTGCTGGCTGACACAGCTGAGCTTCCACAAATCCACGAG 125
 QY 515 SerleuGlutPheHisGluAsnleuGlyHisGlySerPheThrIlePheArgGlySer 534
 Db 126 AGCTGTAGAGTGGCAGAGAACCTGGTCAAGCTCTTTTACCAAGATCTTCGCGGCCG 185
 QY 535 ArgArgGluValaValaAspGlyGluThrHisAspSerGluValaLeuLeuValMetAsp 554
 Db 186 AGGGGGAGGCTGTGGATGTGAGACACATATCTCGGAAGTCCCTCGAAGGTATGAGAC 245
 QY 555 SerArgHisArgAsnCysMetGluSerPheLeuGluAlaAlaSerLeuMetSerGluVal 574
 Db 246 TCCAGACATCGAGACTGATGAGAGCTTTCTTGAAGCCGCAAGCTGATGAGCAAGTA 305
 QY 575 SerTyrProHisLeuValleuLeuHisGlyAlaCysMetAlaGlyAspSerIleMetVal 594
 Db 306 TCCACACCGGACCTGGTGTACTGACGGCTGTCAGAGCTGGAGACAGCATATGAGTG 365
 QY 595 GlnGluPheValTyrLeuGlyAlaAlaIleAspMetTyrLeuArgLysArgGlyHisLeuVal 614
 Db 366 CAGGAATTTTGTATCTAGAGGACATTCATTAACCTGCGCAAGCTGGCCACTGGTG 425

QY 615 SerAlaSerTrpLysLeuGluValaThrLysGlnLeuAlaTyrAlaLeuAsnTyrLeuGlu 634
 Db 426 TCAGCCAGCTGGAAACAGTCAAGGTGACCAACAGCTGCAATATGCCCTTAACCTGAGAG 485
 QY 635 AspLysGlyLeuProHisGlyAsnValSerAlaArgLysValleuLeuAlaArgGly 654
 Db 486 GACAAAGGCTTCTCTCAGGCAACGTCTCAGCAAGAGAGTGCCTCGCTCGAGAGGG 545
 QY 655 GlyAspGlyAsnProPheIleLysLeuSerAspProGlyValSerProThrValleu 674
 Db 546 GGTGATGGAGATCCACCTTTCATTAACTGATGATCTGTGTCACTCCACTGTGCTG 605
 QY 675 SerLeuGluMetLeuThrAspArgGlyLeuProTyrValaAlaProGluCysLeuGluAla 694
 Db 606 AGCTGGAAATGCTCACCAGCAGATACCTGGGTGTGCCCGCAATGTCTCCAGAGGCT 665
 QY 695 GlnThr-LeuCysLeuGluAlaAspLysTrpGlyPheGlyAlaThrThrTrpGluVal 714
 Db 666 CAGCAACTCTGCTTGGAGGCTGACAGTGGGCTTGGAGCCACCACTGGGAAAGTGT 725
 QY 714 eGlnArgGlyProAlaHisIleThrSerLeuGlu-ProAlaLysLysLeuLysPheTyrG 734
 Db 726 CAGCGGGGACCCGCCAAATCACTGCTGGAGGCGCCGCAAAAGCTGAAGCTATG 785
 QY 734 LysArgGlnGlyLysLeuProAlaLeu---LysTrpThrGluLeuAlaGly-LeuIleThr 752
 Db 786 AGGACGAGGAGCAAGTCCCGGCTCTCAATGAGACAGAACTGCCGGAACTATACAC 845
 QY 752 rGlnCys-MetAlaTyrAspProGlyLysArgProSerPheAlaIleLeu-ArgAsp 771
 Db 846 ACACTGATGGCTATGAACTCTGCGCGGCGCCCTTATTCGAGATATCTCTCAGAGAC 905
 QY 772 LeuAsnGlyLeuIleThrSerAspTyrGluLeuLeuSerAspProThrProGlyIlePro 791
 Db 906 CTCACGGGCT-ATTACATTTGTTTACGAGCTGCTCTA-AAACCAACTGGAT---CCC 960
 QY 792 SerProArg 794
 Db 961 AGTCGAGAG 969

RESULT 2
 BC027234
 LOCUS 1625 bp mRNA linear HTC 07-AUG-2002
 DEFINITION Mus musculus, similar to Janus kinase 3, clone IMAGE:3489805, mRNA.
 ACCESSION BC027234
 VERSION BC027234.1 GI:20071020
 KEYWORDS HTC.
 SOURCE house mouse.
 ORGANISM Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 1625)
 DIRECT SUBMISSION
 SUBMITTED (04-APR-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
 NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
 CONTACT: MGC help desk
 EMAIL: cgabs-remail.nih.gov
 TISSUE PROCUREMENT: Lothar Hennighausen Ph.D., Robin Humphreys
 cDNA LIBRARY PREPARATION: Life Technologies, Inc.
 cDNA LIBRARY ARRAYED BY: The I.M.A.G.E. Consortium (LNL)
 DNA SEQUENCING BY: Baylor College of Medicine Human Genome Sequencing Center
 CENTER CODE: BCM-HGSC
 WEB SITE: <http://www.hgsc.bcm.tmc.edu/cdna/>
 CONTACT: amg@bcm.tmc.edu
 GUARANTINEE: P.H., Garcia, A.M., Lu, X., Huijck, S.W., Hale, S.M., Yoon, V.S., Kovis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M., Richards, S., Gibbs, R.A.

REMARK COMMENT

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAC Plate: 35 Row: g Column: 19
 This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein
 This clone has the following problem: Incomplete processing.

FEATURES

Location/Qualifiers
 1..1625

source
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /map="C57BL/6J"
 /clone="IMAGE:3489805"
 /tissue-type="Mammary tumor. MAP-TGF alpha model. 7 months old. gross tissue."
 /clone_lib="NCI_CGAP_Mam5"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"

BASE COUNT 328 a 470 c 483 g 344 t
 ORIGIN

Alignment Scores:

Pred. NO.: 6.07e-127 Length: 1625
 Score: 1358.00 Matches: 281
 Percent Similarity: 75.59% Conservative: 7
 Best Local Similarity: 73.75% Mismatches: 11
 Query Match: 23.17% Indels: 84
 DB: 11 Gaps: 3

US-09-397-967-16 (1-1099) x BC027234 (1-1625)

Oy 1 Metalproproserglutrhproleuileproglinarsercyserserleuenser 20
 Db 42 ATGGACCTCCAGAGGAGACACCTGATCCCTCAGCGCTTTCGACCTTCATCC 101
 Oy 21 SerglualaglyalaLeuHisValLeuLeuProProArgglyProglyProProglinar 40
 Db 102 TCAGAGGAGAGAGCCGATGCTCCTCCCGGAGACGTGGGCTCCCGACGGA 161
 Oy 41 LeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuGlyValAlaAlaAla 60
 Db 162 TTGTCAATCTCTTTGGGAGACTGCTGCGATTTATGCTGCGACTGCCAAGGCC 221
 Oy 61 CysGlyIleLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSer 80
 Db 222 TGTGCACTCCCTGCTTATCATTCGCTTTCGCTGCGACAGAGACTTCTTGC 281
 Oy 81 TrpPheProProSerHisIlePheCysIleGluAspValAspThrGlnValLeuValTyr 100
 Db 282 TGGTTCCCGCCAGACACATCTTGCATAGAGAGAGCTGACATCAAGCTTGGCTAC 341
 Oy 101 ArgLeuAlaGlyPheTyrPheProAspTrpPheGlyLeuGluThrCysHisArgPheGlyLeu 120
 Db 342 AGGCTACGCTTTATTTCCCTGACTGCTTGGGCTGGAGACATGTCACCGCTTTGGGCTG 401
 Oy 121 ArgGlyAspLeuThrSerAlaIleLeuAspLeuHisValLeuGlnHisLeuPheAlaGln 140
 Db 402 CCCAAGATTTACACAGCTCCATCTTGCATGATCTTTAGACAAAGCTTTGGTCCG 461
 Oy 141 HisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLysGluGlnGly 160
 Db 462 CACCCAGAGACCTGCTGAGTGGGCGCCCTCCGCTGGGCTTACATGAAGAGAGAGGGA 521
 Oy 161 GluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGlnGlnAlaGlnArg 180
 Db 522 GAGTTCTCAGGCTGCTGCTGAGCTTGGCCAGATGCTGAGAGAGGCCACGCGC 581
 Oy 181 ProGlyGluLeuLeuLysThrVal 188
 Db 582 CCAGAGAGAGCTGCTGAAGAGCGT - CAGGTGAGAGGCTCGGCGACCCCTTCGTTGCTCTC 640
 Oy 188 ----- 188

Db 641 TTGSGCTGGATCTAGAGCCCTATGCTTTTCCCTTACTGGAACACATACCTCTTCCCA 700
 Oy 188 ----- 188
 Db 701 ACTTGTGAGCACCCCTCTCGTCAGCTGCCCCCGGAAGTCCCTATGTCTGTCT 760
 Oy 189 -----Ser 189
 Db 761 CGTGTCTCTTGGGCCCCCTCACACCCCTAGAGCTGCGGCTGCCCCCCCCCAAGT 820
 Oy 190 TyrIysAlaCysLeuProProSerLeuArgAspValIleGlnGlyGlnAsnPheValThr 209
 Db 821 TACAAAGCCTGCTCCGCCGCCAGCTGCGCATGTGATCCAGGCGCAGAACTTCGTGACA 880
 Oy 210 ArgArgArgIleArgArgThrValValLeuAlaLeuLeuProCysGlyValArgLeuProGly 229
 Db 881 CGCAGGCGCATCCGAGAGCGGTGTGGTCTGGCGTGGC - CGGTGTGTGCTGCTGCGCAGGC 939
 Oy 230 ArgPro-TyrlAlaLeuMetAlaLysTyrIleLeuAspLeuGluArgLeuHisProAlaAl 249.
 Db 940 CGACCGCTACGCGCTCATGCGCCAGATATATTCTGGACCTGGAGCGGCTATCCAGCGGC 999
 Oy 249 aThrThrGluThrPheArgValGlyLeuProGlyAlaGlnGlnGluProGlyLeuLeuArg 269
 Db 1000 CACCAACCGAGACCTTCGCTGTGGGCTCCCGGCGCCCGCAGAGAGAGCCGGGCTTCTGCG 1059
 Oy 269 gValAlaGlyAspAsnGlyIleProTyrSerSerAsnArgIleLeuPheGlnThrPhe 289
 Db 1060 TTGTGGCGGGGAGCAACGCGATCTCTGGAGCTCCGGGAGCCAGC----- 1102
 Oy 289 sasPheProGluIleValAspValSerIle--AsnGlnAlaProArgValGlyProAl 308
 Db 1103 -----GAGGTGCTGTGGGCTGGGCTTGAGAAAGCGGGTCCCGGGGAGCGCTGGC 1152
 Oy 308 a 308
 Db 1153 G 1153

RESULT 3
 LOCUS B1557690 826 bp mRNA linear EST 05-SEP-2001
 DEFINITION 603237050P1 NCI_CGAP_Mam3 Mus musculus cDNA clone IMAGE:3289949 5',
 mRNA sequence.
 ACCESSION B1557690
 VERSION B1557690.1 GI:15445004
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.

REFERENCE NIH-MGC <http://mgc.nci.nih.gov/>.
 1 (bases 1 to 826)
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 CONTACT: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Lotbar Hemmighausen Ph.D., Chu-Xia Deng Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: Incyte Genomics, Inc.

DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: L1AM11733 row: d column: 14
 High quality sequence stop: 760.

FEATURES

source

1..826
 Location/Qualifiers
 /organism="Mus musculus"
 /strain="129,C57BL/6J,FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:5289949"
 /clone_lib="NCI_CGAP_Mam3"

```

/tissue_type="tumor, gross tissue"
/dev_stage="10 months"
/lab_host="DH10B"
/Note="Organ: mammary; Vector: pCMV-Sport6; Site:1: Salt;
Site:2: NotI; Cloned unidirectionally. Primer: oligo dT.
Library constructed by Life Technologies. Investigators
providing samples: Lohar Hemighausen/Chu-Xia Deng, NIH
Reference for transgenic model: Xu et al., Nature Genetics
22, 37-43 (1999)."
```

BASE COUNT 147 a 252 c 240 g 187 t

ORIGIN

Alignment Scores:

Pred. NO.:	1,58e-110	Length:	826
Score:	1192.00	Matches:	249
Percent Similarity:	95.45%	Conservative:	3
Best Local Similarity:	94.32%	Mismatches:	9
Query Match:	20.34%	Indels:	7
DB:	13	Gaps:	1

US-09-397-967-16 (1-1099) x B1557690 (1-826)

```

QY      1 MetAlaProProSerGluGluThrProLeuIleProGlnArgSerCysSerLeuSerSer 20
         |||||||
Db      37 ATGGACCTCCAGAGTGAAGACACCTGTGATCCCTCAGCGCTCTCAGCTCTCACC 96
QY      21 SerGluAlaGlyAlaLeuHisValLeuLeuProProArgGlyProGlyProProGlnArg 40
         |||||||
Db      97 TCAGAGCAGAGAGCCCTGCATGTGCTCTTCCTCCCGGGGACCTGGGCTCCCGCCACGGA 156
QY      41 LeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuValArgAlaAlaAlaAlaAla 60
         |||||||
Db      157 TTGTGATTCCTC-TTGGGGAGCTACTGCTGGAGCATTTATGTGCGAGCGCCAGGCC 215
QY      61 CysGlyIleLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSerCys 80
         |||||||
Db      216 TGTGGCATCTGCTGCTTTTATCATCTGCTTTGCTGCTGCGCCTAGAGCTTCTTTC 275
QY      81 TrpPheProProSerHisIlePheCysIleGluAspValAspThrGlnValLeuValTyr 100
         |||||||
Db      276 TGGTTTCCCAAGCCACATCTTCTGCATAGAGACCTGGACACTCAAGCTTGTGCTAC 335
QY      101 ArgLeuArgPheTyrPheProAspTyrPheGlyLeuGluThrCysHisArgPheGlyLeu 120
         |||||||
Db      336 AGGCTAGCTTTTATTTCCCTGACTGCTTGGCTGGAGCATGTCAACGCTT-GGGCTG 394
QY      121 ArgLysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLeuPheAlaGln 140
         |||||||
Db      395 CGCAAGACATTG-ACCACTGCCATCCTTGACTTACATGTTTAGACATCTCTTGTCTCAG 453
QY      141 HisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLysGluGlnGly 160
         |||||||
Db      454 CACCGGAGTACCTGGTGGAGTGGGGCCCTCCGGTGGCCTTAGCATGAAGAGACAGGA 513
QY      161 GluPheLeuSerLeuAlaValIleuAspLeuAlaGlnMetAlaArgGluGlnAlaGlnArg 180
         |||||||
Db      514 GAGTCTCTGAGCTGGCGGCTGGAGACTTGGCCAGATGCTGCTGAGCAGGCCACAGCC 573
QY      181 ProGluGluLeuLeuLeuTyrValSerTyrLysAlaGlyLeuProProSerLeuArgasp 200
         |||||||
Db      574 CCAGGAGAGTGTGAGAGAGGCTGAGTACAAAGCCGTGTGCGCCACCTGCGGGAT 633
QY      201 ValIleGlnGlyLnsnspPheValThrArgArgArg-1leArgArgThrValValLeuAl 220
         |||||||
Db      634 GTATATCCAGGGCCAGAACTTCGTGACACGACGACGACATCCGACGAGACCTGCTTGGC 693
QY      220 AlaLeuProCysGlyArgLeuProGlyArgProTyrAlaLeuMetAlaLysTyrIleLe 240
         |||||||
Db      694 GCTGCG-CCGTGTGTGCTGCTTGCAGGCGACCGTACGCTCAGCCCAAGTATATTCT 752
QY      240 uAspLeuGluArgLeuHisPro-AlaAlaThrThrGluThrPheArgValGlyLeuProG 260
         |||||||
Db      753 GGACCTGGAAACGTTACATCCAGACGCAACAGAGA--CTCCGTGTGTGGGCTCCCGG 809
```

```

QY      260 lYalagln 262
         |||||||
Db      810 CGCCGACAG 817
RESULT 4
BM935182
LOCUS
DEFINITION
BM935182
698 bp mRNA linear EST 13-MAR-2002
UI-M-BH3-ag2-f-08-0-UI.r1 NIH_BMAP_M.S4 Mus musculus cDNA clone
UI-M-BH3-ag2-f-08-0-UI 5', mRNA sequence.
ACCESSION
BM935182
VERSION
BM935182.1 GI:19394334
KEYWORDS
EST.
SOURCE
house mouse.
ORGANISM
Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 698)
Bonaldo,M.F., Lennon,G. and Soares,M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
JOURNAL
MEDLINE
97044477
COMMENT
Contact: Chin, H
National Institute of Mental Health
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
20892-9643, USA
Tel: 301 443 1706
Fax: 301 443 9890
Email: MEST@mail.nih.gov
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
Seq primer: M13 REVERSE
FEATURES
Source
Location/Qualifiers
1..698
/organism="Mus musculus"
/strain="C57Bl/6J"
/db_xref="taxon:10090"
/clone="UI-M-BH3-ag2-f-08-0-UI"
/clone_11b="NIH_BMAP_M.S4"
/dev_stage="27-32 days"
/lab_host="DH10B (Life Technologies)"
/Note="Vector: pT7AD-Pac (Pharmacia) with a modified
polylinker; Site:1: Not I; Site:2: Eco RI; The
NIH_BMAP_M.S4 library is a subtracted library of a series,
ultimately derived from a mixture of individually tagged
normalized libraries from ten regions of the mouse brain
(cerebellum, brain stems, olfactory bulbs, hypothalamus,
cortex, amygdala, basal ganglia, pineal gland, striatum,
hippocampus) after a series of subtractions to reduce the
representation of cDNAs from which ESTs had already been
generated. The following serially subtracted libraries
were generated in this process: NIH_BMAP_M.S4,
NIH_BMAP_M.S3.3, NIH_BMAP_M.S3.2, NIH_BMAP_M.S3.1,
NIH_BMAP_M.S2, NIH_BMAP_M.S1. The subtracted library
(NIH_BMAP_M.S4) was constructed as follows: PCR amplified
cDNA inserts from NIH_BMAP_M.S3.3, NIH_BMAP_M.S3.2, and
NIH_BMAP_M.S3.1 clones from which 3' ESTs had been derived
was used as a driver in a hybridization with a pool of
the NIH_BMAP_M.S3.3, NIH_BMAP_M.S3.2, and NIH_BMAP_M.S3.1
libraries in the form of single-stranded circles. The
remaining single-stranded circles (subtracted library)
was purified by hydroxyapatite column chromatography,
converted to double-stranded circles and electroporated
into DH10B bacteria (Life Technologies) to generate the
NIH_BMAP_M.S4 library. This procedure has been previously
described (Bonaldo, Lennon and Soares, Genome Research
6:791-806, 1996)"
BASE COUNT 132 a 211 c 215 g 138 t 2 others
ORIGIN
```

Alignment Scores:

Pred. No.:	1,88e-105	Length:	698
Score:	1141.00	Matches:	225
Percent Similarity:	96.15%	Conservative:	0
Best Local Similarity:	96.15%	Mismatches:	6
Query Match:	19.47%	Indels:	4
DB:	14	Gaps:	2

US-09-397-967-16 (1-1099) x BM935182 (1-698)

```

OY 792 SerProArGAspGluLeuGluCysValAlaGluGluLeuTyrAlaCysGlnAspProAla 811
    |||||||
DB 3 AGTCCCTCGAGATGAGCTGTGTC---GGTGGCGCCGAGCTCTATGCTGCGCCAGAGACCCCGCC 59
    |||||||

OY 812 IllePheGluGluArgHisLeuLysTyrTrpSerLeuLeuGlyLysGlnAsnGlySer 831
    |||||||
DB 60 AATATTCGAGGAGAGACACCTTAAGTACATCTCTTGGTGGCAAGGCAACTTTGGCAGC 119
    |||||||

OY 832 ValGluLeuGluCysArgTyrAspProLeuGluYAspAsnThrGlyProLeuValAlaValLys 851
    |||||||
DB 120 GTGGAGCTGTGCGCTGATGACCCCTGGGGGACATACGGGACCCCTGGTGGCAGTGAA 179
    |||||||

OY 852 GlnLeuGlnHisSerValProAspGlnGlnArgAspPheGlnArgGluIleGlnIleLeu 871
    |||||||
DB 180 CAGCTACAGACACAGCGGGCCAGACAGAGGAGACTTCAGCGGAGATTCAGATCCTT 239
    |||||||

OY 872 LysAlaLeuHisSerAspPheIleValLysTyrArgGlyValSerTyrGlyProGlyArg 891
    |||||||
DB 240 AAGGCTGTGACAGGACATTCATGTCAGTACCGGGAGTACAGCTATGCGGACAGTTCGC 299
    |||||||

OY 892 GlnSerLeuArgLeuValMetGluTyrLeuProSerGlyCysLeuArgAspLeuGln 911
    |||||||
DB 300 CAGACCTGCGGTGTGTGATGAGTACCTGCCAGCGGCTGCTGCGAATCTTCGCGAG 359
    |||||||

OY 912 ArgHisArgGly---LeuHisThrAspArgLeuLeuPheAlaTrpGlnIleCysLys 930
    |||||||
DB 360 CCCCATCGCGCGCCGCTGCACACCGCCCTACTGCTGTTGCTTGGCGCAGATCTGCAAG 419
    |||||||

OY 931 GlyMetGluTyrLeuGlnValArgArgCysValHisArgAspLeuAlaIleArgAsnIle 950
    |||||||
DB 420 GGCATGTGACTACCTGGGTGGCGCGCCGCTGCTACACCGTACCTGGCTGGCGCAACATC 479
    |||||||

OY 951 LeuValGluSerGluAlaHisValLysIleAlaAspPheGlyLeuAlaLysLeuLeuPro 970
    |||||||
DB 480 TTGGTGGAGACGAGGCTCATGTGAAGATCGCGGACTTGGCTGCTGAAGCTGCTGCGCC 539
    |||||||

OY 971 LeuGlyLysAspTyrTyrValValArgGluProGlyGlnSerProIle-PheTrpTyrAl 990
    |||||||
DB 540 CTGGGAAAGGACTACTACTGTGCTGGCGGAGCTGGCCAAAGCCCAATCTTNTTGGTATGC 599
    |||||||

OY 990 AProGluSerLeuSerAspAsnIlePheSerArgGlnSerAspValTrpSerPheGlyVal 1010
    |||||||
DB 600 CCCGAGATCTCTATCTGACAC-ATCTTCTCNCGCCAATGTGACGTGGAGCTTCGGAAT 658
    |||||||

OY 1010 ValLeuTyrGluLeuPheThrTyrCysAspLysSerCys 1023
    |||||||
DB 659 GGTGTGTGACGAGCTCTTCACTACTGCGACAAGAGCTGC 698
    |||||||

RESULT 5
LOCUS Bf150250 675 bp mRNA linear EST 29-DEC-2000
DEFINITION uy83e09.y1 NCI_CGAP_Mam5 Mus musculus cDNA clone IMAGE:3666184 5'
similar to TR:p97423 p97423 JANUS KINASE 3 ;, mRNA sequence.
ACCESSION Bf150250
VERSION Bf150250.1 GI:11031645
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 675)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  
```

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Other-ESTs: uy83e09.x1
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Lotmar Hennighausen Ph.D., Robin Humphreys
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Washington University Genome Sequencing Center
Clone Distribution: NCI-CGAP clone distribution Information can be
found through the I.M.A.G.E. Consortium/LLNL at:
image.llnl.gov/image/html/iresources.shtml
MGI:1426952
Seq primer: -40RP from GIBCO
High quality sequence stop: 390.
Location/Qualifiers
1..675
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:3666184"
/clone_11b="NCI_CGAP_Mam5"
/tissue_type="tumor, gross tissue"
/dev_stage="7 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-Sport6; Site: 1; Salt:
Site: 2; Note: Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigators
providing samples: Lotmar Hennighausen/Robin Humphreys,
NIH"

BASE COUNT 138 a 215 c 197 g 123 t 2 others
ORIGIN

Alignment Scores:

Pred. No.:	9,15e-105	Length:	675
Score:	1134.00 <td>Matches:</td> <td>217 </td>	Matches:	217
Percent Similarity:	96.02% <td>Conservative:</td> <td>0 </td>	Conservative:	0
Best Local Similarity:	96.02% <td>Mismatches:</td> <td>7 </td>	Mismatches:	7
Query Match:	19.35% <td>Indels:</td> <td>2 </td>	Indels:	2
DB:	12	Gaps:	2

US-09-397-967-16 (1-1099) x Bf150250 (1-675)

```

OY 705 GlyPheGluAlaThrThrTrpGluValPheGlnArgGlyProAlaHisIleThrSerLeu 724
    |||||||
DB 1 GGCTTTGGAGCCACACAGCTGGAGGTGTTTCAGCGGGGACCCGCCACATCAGCTGCTG 60
    |||||||

OY 725 GluProAlaLysLysLeuLysPheTyrGluAspGlnGlyGlnLeuProAlaLeuLysTrp 744
    |||||||
DB 61 GAGCCCGCCAAAAGCTGAAGTTCATGAGGACACGAGACAGCTGCCGCTCAAAATGG 120
    |||||||

OY 745 ThrGluLeuAlaGlyLeuIleThrGlnCysMetAlaTyrAspProGlyArgArgProSer 764
    |||||||
DB 121 ACAGAACTGCGCGGACTTTCACACAGTCAATGCGCTATGATCCCGCGGCCCTCTCC 180
    |||||||

OY 765 PheArgAlaIleLeuArgAspLeuAsnGlyLeuIleThrSerAspTyrGluLeuLeuSer 784
    |||||||
DB 181 TTCCGAGCTATCTCAGAACTCTCAACGCGCTTCATTACATCAGATTACAGACTCTCTCA 240
    |||||||

OY 785 AspProThrProGlyIleProSerProArGAspGluLeuGluCysValAlaGluAlaGlnLeu 804
    |||||||
DB 241 GACCCACACACTGGAGATCCCGAGTCTCGATGATGACTGTGTC---GGTGGCGCCAGCTC 297
    |||||||

OY 805 TyrAlaCysGlnAspProAlaIlePheGluGluArgHisLeuLysTyrTrpSerLeuLeu 824
    |||||||
DB 298 TATGCTGTCCAGACCCGCCCATATTCGAGGAGAGACACCTTAAGTACATCTCTTGGCTG 357
    |||||||

OY 825 GlyLysGlyAsnPheGlySerValGluLeuGluCysArgTyrAspProLeuGluYAspAsnThr 844
    |||||||
DB 358 GGCMAAGGCACTTTGGCAGCTGTGGAGCTGTGCGCTATGACCCCTGGGGGACATATACG 417
    |||||||
  
```

QY 845 G1yProLeuValAlaValLysGlnLeuGlnHisSerValProAspGlnGlnArgAspPhe 864
 |||
 Db 418 GGACCCCTGGTGGAGTGAACACAGCTACAGCAGCGGCGCACACAGCAGAGGACTTC 477
 QY 865 G1aArgL1u1eGln1leuLysAlaLeuHisSerAspPheLeuValLysTyrArgLy 884
 |||
 Db 478 CAGGGAGATCCAGATCTTAAGCTCTGACAGCGACTTCACTGCAAGTACCGGGGA 537
 QY 885 ValSerTyrG1yProG1yArgGlnSerLeuArgLeuValMetG1uTyrLeuProSerG1y 904
 |||
 Db 538 GTGAGCTATGGGCGCANTGCCAGAGCCTGGGCTTGATGATGAGTACCTGCCAGCGGC 597
 QY 905 CysLeuArgAspLeuLeuGlnArgHisArgLy---LeuHisTyrAspArgLeuLeu 923
 |||
 Db 598 TGGCTGGAGACTTCTCAGAGCGCCATCGCGCGCTGCACACCGAGCTACTCTG 657
 QY 924 PheAlaTyrGlnIleCys 929
 |||
 Db 658 TTGCCTGGCAGATCTGC 675
 RESULT 6
 BG870513 907 bp mRNA linear EST 29-MAY-2001
 LOCUS 6027914328F1_NCL_CGAP_SG2 Mus musculus cDNA clone IMAGE:4922742 5',
 DEFINITION mRNA sequence.
 ACCESSION BG870513 GI:14221053
 VERSION BG870513.1
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus.
 Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathii; Muridae; Murinae; Mus.
 1 (bases 1 to 907)
 NIH-MGC http://mgc.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.
 CDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: L1AM10842 row: P column: 07
 High quality sequence stop: 742.
 FEATURES
 source Location/Qualifiers
 1..907
 /organism="Mus musculus"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:4922742"
 /clone_lib="NCL CGAP_SG2"
 /lab_host="DH10B (T1 phage-resistant)"
 /note="Organ: salivary gland; Vector: pCMV-SPORT6; Site: 1;
 NotI; Site: 2; SalI; Cloned unidirectionally. Primer: Oligo
 dT. Average insert size 1.3 kb. Constructed by Life
 Technologies. Note: this is a NCL CGAP Library."
 BASE COUNT 166 a 279 c 275 g 185 t 2 others
 ORIGIN
 Alignment Scores: Length: 907
 Pred. No.: 1031.50 Matches: 216
 Score: 1031.50
 Percent Similarity: 88.93% Conservative: 1
 Best Local Similarity: 88.52% Mismatches: 23
 Query Match: 17.608 Indels: 7
 DB: 12 Gaps: 3
 US-09-397-967-16 (1-1099) x BG870513 (1-907)
 QY 858 ProAspGlnGlnArgAspPheGlnArgGlnL1eGln1leuLysAlaLeuHisSerAsp 877

|||
 Db 39 CCAGACAGAGAGGAGCTTCACAGCGGAGATTCAGATCTTAAGGCTTCACAGCGAC 98
 QY 878 Phe1leValLysTyrArgG1yValSerTyrG1yProG1yArgGlnSerLeuArgLeuVal 897
 |||
 Db 99 TTCATGTCAGATCCAGGAGGAGTCACTATGGCCAGAGTCCACAGAGCTTGGTGGTG 158
 QY 898 MetG1uTyrLeuProSerG1yCysLeuArgAspLeuLeuGlnArgHisArgLy---Leu 916
 |||
 Db 159 ATGAGTACCTGCCAGGCGGTGCTGCGAGACATTCCTGCAGCGCCATCCGCGCGCTG 218
 QY 917 HisThrAspArgLeuLeuPheAlaTyrGlnIleCysLysG1yMetG1uTyrLeuGly 936
 |||
 Db 219 CACACCGACCGCTTACCTGCTTCTGCGAGATCTTCGACAGGCGCAGTACTCTGAGT 278
 QY 937 AlaArgArgCysValHisArgAspLeuAlaAlaArgAsn1leuValGlnSerG1uAla 956
 |||
 Db 279 GCGGCGCGCTGCTGACACCGAGACCTGCTGCGCACATCTTGTTGAGAGGAGGCT 338
 QY 957 HisValLys1leAlaAspPheG1yLeuAlaLysLeuLeuProLeuG1yLysAspTyr 976
 |||
 Db 339 CATGTGAAGATCGCGGAGCTTGGCTCGTAACTGCTGCCCTGGGGAAGAGTACTAC 398
 QY 977 ValValArgG1uProG1yGlnSerProIlePheTyrAlaProG1uSerLeuSerAsp 996
 |||
 Db 399 GTGGTCGCGAGCTGCGCAAGGCCCATCTTGTGATGCCCCGAGAGTCTATCTGAC 458
 QY 997 Asn1lePheSerArgGlnSerAspValTyrSerPheG1yValLysLeuTyrG1uLeuPhe 1016
 |||
 Db 459 AACATCTTCTCCGCCATCTGAGACTGTGAGCTTGGAGTGGTGTGTGAGACTCTTC 518
 QY 1017 ThrTyrCysAspLysSerCysSerProSerAlaGlnPheLeuArgMetMetG1yProG1u 1036
 |||
 Db 519 AC-TACTGCGACAGAGCTGAGNCCATCCGCTAGTTCCTGAGCAGATGATGGGCGTAG 577
 QY 1037 ArgGlnG1yProProLeuCysArgLeuGlnLeuLeuAlaGlnG1yArgArgLeuPro 1056
 |||
 Db 578 CGTGAAGAGACCCCGCTTCTGCGCTC-CTGAGAGCTGTGCGAGGGGCCGAGC-CTCCCA 635
 QY 1057 ProProProThrCysProThrGlnValGlnGlnLeuMetGlnLeuCysTyrPalaProG1u 1076
 |||
 Db 636 CATCTTCCACTGCG---CACGAGGTTCAAGAGCTCATGACGCTGTCTGGGCCCCAGC 692
 QY 1077 ProHisAspArgProAlaPheAlaThrLeuSerProGlnLeuAspProLeuTyrArgLy 1096
 |||
 Db 693 CCG---ACGAGCGGCGAGCTTGGAGCCCTGAGCCCGACGCTGAGCGGCTGTGAGTGA 749
 QY 1097 -ArgProGly 1099
 |||
 Db 750 AAGACCCGGG 759
 RESULT 7
 BG873355 720 bp mRNA linear EST 29-MAY-2001
 LOCUS 602794328F1_NCL_CGAP_SG2 Mus musculus cDNA clone IMAGE:4925598 5',
 DEFINITION mRNA sequence.
 ACCESSION BG873355
 VERSION BG873355.1 GI:14223895
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus.
 Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathii; Muridae; Murinae; Mus.
 1 (bases 1 to 720)
 NIH-MGC http://mgc.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.
 CDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at:

http://image.lnl.gov

Plate: L1A010850 row: 9 column: 07

High quality sequence stop: 709.

FEATURES

Location/Qualifiers

1..720

/organism="Mus musculus"

/strain="FVB/N"

/db_xref="taxon:10090"

/clone="IMAGE:4925598"

/lab_host="NCI-CGAP-SG2"

/note="Organ: salivary gland; Vector: PCMV-SPORE6; Site: 1;

Note: Site 2: Salt; Cloned unidirectionally. Primer: Oligo

dr. Average insert size 1.3 kb. Constructed by Life

Technologies. Note: this is a NCI-CGAP Library."

BASE COUNT

123 a 229 c 229 g 129 t

ORIGIN

Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
Score:	5,81e-94	720	201	0	8	1
Percent Similarity:	1028.50					
Best Local Similarity:	95.71%					
Query Match:	17.55%					
DB:	12					

US-09-397-967-16 (1-1099) x BG873355 (1-720)

891 ArgGlnSerLeuArgLeuValMetGluTyrLeuProSerGlyCysLeuArgAspLeu 910

1 CGCCAGAGCTGGGCTGGTGGATGAGTACCTCCAGCGGCTGCGAGACTTCCTG 60

911 G1AArgH1sArgGly---LeuH1sThAspArgLeuLeuLeuLeuLeuLeuLeu 929

61 CACGGCAGTGGGCGGCGCGACACCGACCGGCTGCTGCTGCTGCTGCTGCTGCTG 120

930 LysGlyMetC1uTyrLeuGlyAlaArgArgCysValH1sArgAspLeuAlaArgAsn 949

121 AAGGGATGAGTACCTGGCTGGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180

950 IleLeuValGlnSerGlnAlaH1sValysIleAlaAspPheGlyLeuAlaLysLeu 969

181 ATCTTGTGAGAGCGAGGCTCTGTGAAGATCGCGACTTCGCTGCTGCTGCTGCTG 240

970 ProLeuGlyLysAspTyrTyrValAlaArgGluProGlyGlnSerProIlePheTyr 989

241 CCCCTGGAAAGACTACTAGCTGCTGCGGAGCTGCGCAAGCCCATCTTTGCTAT 300

990 AlaProGlnSerLeuSerAspAsnIlePheSerArgGlnSerAspValIrrSerPheGly 1009

301 GCCCGGAGTCCCTATCTACACATCTCTCCGCCAATCTGACCTGTGAGACTTCGGA 360

1010 ValValLeuTyrGlnLeuPheTyrTyrCysAspLysSerCysSerProSerAlaGluPhe 1029

361 GTGGCTGTGACAGCTTCTACACTCTGCGACAGAGCTGACGCCATCCGCTGAGTTC 420

1030 LeuArgMetMetC1yProGluArgGluGlyProProLeuGlnCysArgLeuGluLeu 1049

421 CTGGCGATGATGGGCTGAGGCTGAGAGACCCCGCTGCGGCTCTGAGAGCTGCTG 480

1050 AlaGlnGlyArgArgLeuProProProProProProProProProProProProPro 1069

481 GCAGAGGCGCGAGGCTCCACCACTCTCCACCTGCGCGAGGTTTACGAGAGCTCATG 540

1070 GlnLeuCysTrrPalaProGluProH1sAspArgProAlaPheAlaThrLeuSerProGln 1089

541 CAGCTGTGCTGGGCGCGCGCGCA-GACCGGCGAGCTTGGG-ACCCTGAGCCGCCAG 598

1090 LeuAspProLeuTrrPargGlyArgProGly 1099

599 CTGGAGCGCTGTGGCTGAGAGACCCGGA 628

RESULT 8

AW742526

LOCUS

DEFINITION

AW742526

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgabbs-remail.nhl.gov

This clone is available royalty-free through LINL; contact the

IMAGE Consortium (info@image.lnl.gov) for further information.

Seq primer: -40RP from Gibco

High quality sequence stop: 471.

Location/Qualifiers

1..621

/organism="Mus musculus"

/strain="C3H x 101 F1"

/db_xref="taxon:10090"

/clone="IMAGE:2780302"

/clone_lib="Soares_mouse_NMIE"

/sex="male"

/dev_stage="newborn"

/lab_host="DH10B"

/note="Organ: inner ear, 170 pooled; Vector: pRT3D-Pac;

Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed

with a Not I - oligo(dT) primer (5'

TGTTCACATCTGACAGTGGAGGCGCGCCACCTTTTCTTTTCTTTTCTTTTCTTTT

3'); double-stranded cDNA was ligated to Eco RI adaptors

(Pharmacia), digested with Not I and cloned into the Not

I and Eco RI sites of the modified pRT3D vector. Library

is normalized, and was constructed and donated by Bento

Soares and M. Fatima Bonaldi (University of Iowa) and R.

Hardisty, A. Varela-Carver, P. Mburu and S.D.M. Brown (MRC

UK Mouse Genome Centre and Mammalian Genetics Unit,

Harwell, UK)."

BASE COUNT

124 a 197 c 182 g 118 t

ORIGIN

Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
Score:	8.47e-93	621	196	1	8	2
Percent Similarity:	1016.00					
Best Local Similarity:	95.17%					
Query Match:	94.69%					
DB:	17.34%					
DB:	10					

US-09-397-967-16 (1-1099) x AW742526 (1-621)

741 AlaLeuLysTrrPthGlnLeuAlaGlyLeuIleThrGlnCysMetAlaTyrAspProGly 760

3 GCTCTATATGACAGACTGCGGAGCTTATGACACAGCTCATGATGATGATGCTGCG 62

761 ArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIleThrSerAspTyr 780

63 CGGCGCCCTCTCTCGAGCTATCTCAGAGACCTCAACGGCTCTATTAATCAAGATTAC 122

781 GlnLeuLeuSerAspProThrProGlyTyrLeuProSerProAlaGluLeuCysValAla 800

123 GAGCTCTCTCAGACCCACACACCTGCGCATGCCGAGTCTCGAGATGAGACTGTC---GCT 179

801 GlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGluGluArgH1sLeuLysTyr 820

```

|||||
Db 180 GGGCCAGCTATGCTCTCCAGACCCGCCATATTCAGAGAGACACCTTAAGTAC 239
Oy 821 11SerLeuLeuGlyLysGlyAsnPhelGlySerValGluLeuGlyAspArgTyraProLeu 840
Db 240 ATCTCTTGTGCTGGCAGGGCACTTTGGCAGCGTGAAGTGTGCCCTATGACCCCTTG 239
Oy 841 GlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGlnHisSerValProAspGln 860
Db 300 GGGGACATACGGGACCCCTGTGGCAGTGAACACCTACAGCACAGCGGGCCAGACAG 359
Oy 861 GlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuHisSerAspPheIleVal 880
Db 360 CAGAGGAGCTTCCAGCCGGAGATTACATCTTAAGCTTCGACAGCAGCTTATGCTC 419
Oy 881 LysTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuArgLeuValMetGluTyr 900
Db 420 AAGTACCGGGAGTACAGTATGGCCAAAGTCCAGAGCCCTGCGTTGTGATGAGTAC 479
Oy 901 LeuProSerGlyCysLeuArgAspLeuLeuGlnArgHisArgGly---LeuHisThrAsp 919
Db 480 CTGCCACGGCTCTCGTGGAGACTCTCTGACGCGCATCGCGCGCTTSCACACCGAC 539
Oy 920 ArgLeuLeuPheAlaTrpGlnIleCysLysGlyMetGluTyrLeuGlyAlaArgArg 939
Db 540 CGGCTACTGCTGTCGCTGGCAGATCGCAAGGCAATGAGTACTGTGTGCGCGCGC 599
Oy 940 CysValHisArgAspLeuAla 946
Db 600 TGCCTACACCTGACCTGGCT 620

RESULT 9
AA920299 731 bp mRNA linear EST 20-APR-1998
LOCUS AA920299
DEFINITION v93g02.1 Soares_thymus_2nbmt Mus musculus cDNA clone
IMAGE:1282802 5' similar to gb:L40172 Mus musculus JAK3 gene,
complete cds (MOUSE);, mRNA sequence.
AA920299
AA920299.1 GI:3067078
EST.
SOURCE house mouse.
ORGANISM Mus musculus.
REFERENCE 1 (bases 1 to 731)
AUTHORS Marra,M., Kucaba,T., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisels,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
TITLE The WashU-HM Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HM Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Seq primer: -28m13 rev2 ET from Amersham
High quality sequence stop: 215.
FEATURES
Source
1..731
location/Qualifiers
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:1282802"
/clone_1id="Soares_thymus_2nbmt"
/sex="male"
/tissue_type="Thymus"

```

```

/dev_stage="4 weeks"
/lab_host="DH10B"
/vector="pUT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was primed with a Not I-oligo(dT) primer [5',
TGTACCAATCTGTAAGTGGAGCGGCCGGTGTATTTTATTTTATTTT
3']; double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pUT73 vector. RNA
provided by Dr. Bertrand Jordan. Library went through two
rounds of normalization, and was constructed by Bento
Soares and M. Fatima Bonaldo."
BASE COUNT 167 a 209 c 214 g 141 t
ORIGIN
Alignment Scores:
Pred. No.: 2,87e-92 Length: 731
Score: 1012.00 Matches: 213
Percent Similarity: 90.42% Conservative: 4
Best Local Similarity: 88.75% Mismatches: 21
Query Match: 17.27% Indels: 1
DB: Gaps: 6

US-09-397-967-16 (1-1099) x AA920299 (1-731)
Oy 660 ProPheIleLysLeuSerAspProGlyValSerProThrValLeuSerLeuGluMetLeu 679
Db 9 CCTTTCATTAAGCTGAGTGAATCCGTGTCAGTCCACTGCTGAGCCCGAAATGCTC 68
Oy 680 ThrAspArgIleProTrpValAlaProGlyCysLeuGlnGluAlaGlnThrLysLeu 699
Db 69 ACCACAGAAATACCTCTGGTGGGCCCGCAATGCTCCAGAGAGCTCAGACACTCGGCTTG 128
Oy 700 GlnAlaAspLysTrpGlyPheGlyAlaThrThrTrpGluValPheGlnArgGlyProAla 719
Db 129 GAGGCTGACAAAGTGGGCTTTGGAGCCACACGAGGAGGTGTACACGGGGGCCCGC-- 186
Oy 720 HisIleThrSerLeuGluProAlaLysLysLysSerPheTyrGluAspGlnGlyGlnLeu 739
Db 187 CACATCACCTCGCTGGAGCCCGCCAAAAGCTGAAGTTTATGAGGACGACGACACTG 246
Oy 740 ProAlaLeuLysTrpThrGluLeuAlaGlyLeuIleThrGlnCysMetAlaTyrAspPro 759
Db 247 CCCGCTCTCAATATGACAGAACTGGCGGACTTTCACAGAGTCAGCGCTATGATTC 306
Oy 760 GlyArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIleThrSerAsp 779
Db 307 GCGCGGCGCCCTCTCCGAGTAACTCAGACACCTCAACGGCCATTTACATCAGAT 366
Oy 780 TyrGluLeuLeuSerAspProThrProGlyIleProSerProAlaAspGluLeuCysVal 799
Db 367 TACGAGCTCTCTCAGACCCCAAAACTGCGATCCCGAGTCTCGAGATGAGACTGTGC-- 423
Oy 800 AlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGluGluArgHisLeuLys 819
Db 424 GGTGGGGCCGAGCTCTATGCTCGCAGAGACCCCGCAATATTCAGAGAGACACACTTAAG 483
Oy 820 TyrIleSerLeuLeuGlyLysGlyAsnPhelGlySerValGluLeuGlyAspArgTyraPro 839
Db 484 TACATCTCTTTTGTCT-GGGAGAGGGAACCTTGGCAGCGTGTGAGCTGTGCGCTATGACCCC 542
Oy 840 LeuGlyAspAsnThrGlyProLeuValAlaValLys-GlnLeuGlnHisSerValProAs 859
Db 543 CTGGGAGACAATACCGGGACCCCTGTGGCAGTGAACACAGCTACAGCAGCGGGGCGAGA 602
Oy 859 pGlnGlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuHisSerAspPheI 879
Db 603 CCAGCAAAAGAGACTTCCAAAGCGAGATTCAGATTCCTTAAGCTCTGCACACGAGATTCTAT 662
Oy 879 eValLysTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuArgLeuValMet 898
Db 663 GCTCAAGTACCGGGGAATCAG-TATGGGCGAGGTGCCAAGC-CTCGGTTTGATATG 718

```

TITLE
JOURNAL
COMMENT

GAGAGAGAGAGATCCACAGAGCTCTTTTATTTTTTTTTTNN 3'). cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 185.0. Second strand cDNA was prepared with the primer adapter of sequence (5' GAGAGAGAGATTCTCGAGATTATTAATATCCCTCCCTCC 3'). cDNA was cleaved with XhoI and BamHI. Vector: a modified pbluescript KS(+) after bulk excision from Lambda FLC I.	
BASE COUNT	106 a 171 c 156 g 150 t
ORIGIN	
Alignment Scores:	
Pred. No.:	3.88e-07 length: 583
Score:	960.00 Matches: 183
Percent Similarity:	98.94%
Best Local Similarity:	97.34%
Query Match:	15.38%
DB:	
	10
	Gaps: 0.

[illegible]

Oy	223	ProCysgIyArLeuProGlyArPro-TyrAlaLeuMetAlaLysTrIleLeuSspIle	242
Db	15	CGCGCGGCGCGCTCCAGCGAGACCGCGCATCTCATGCGCAAGTACATGAGCCT	74
Oy	242	uGluArGLeuHISProAlaAlaTrhThGluThPheArGValGlyLeuProGlyAla--	261
Db	75	GGAGCGGCTGGATTCAGCGGGGGCGCGAGACCTTCACGTGGGCGCTTCCGTGGGCGCT	134
Oy	262	-----GInGluGluProGlyLeuLeuArGValAlaGlyAspAsnGlyIleProTrpSe	279
	:::		
Db	135	TGCTGGCCACGACGGGCTGGGCTGCTCCGGTGGCTGGTACGCGCGCATCGGCTGGAC	194
Oy	279	rSerAsnAsp----GluLeuPheGlnThPheCysAspPheProGluIleValAspAlse	298
	:::		
Db	195	CCAGGAGAACGAGGAGGCTCCACGCCCTTCTGGCACCTTCCAGAAATCGTAGACATTAG	255
Oy	298	rIleAsnGlnAlaProArGValGlyProAlaGlyLuhHisArGLeuValThrValThrAr	318
Db	255	CATCAAGCAGGCGCGCGGCTGGCGCGCGGCGGAGACCGCGCTGGTCTTACAG	314
Oy	318	gMeAspGlyHisIleLeuGluAlaGluPheProGlyLeuProGluAlaLeuSerPheVa	338
Db	315	GACAGACAAACAAATTTTGAGGCGGAGTTCCAGAGGCTGCCCGAGGCTGTGGTTCGT	374
Oy	338	lAlaLeuValAspGlyTyTrPheArGLeuIleCysAspSerArGHisTyTrPheCysLysG	355
Db	375	GGCGCTGGTGAGCGGCTACTTCGGGTGACACAGACATCCACAGCACTTCTTCGTGAAGA	433

BASE COUNT	106 a	171 c	160 g	118 t	4 o
ORIGIN					
Alignment Scores:					
Pfed. No.:	4.13e-84		Length:		559

Score: 930.00 Matches: 179
 Percent Similarity: 96.24% Conservative: 0
 Best Local Similarity: 96.24% Mismatches: 7
 Query Match: 15.87% Indels: 1
 DB: 12 Gaps: 0

US-09-397-967-16 (1-1099) x BG800853 (1-559)

OY 415 ThrProLeuGlyProAspTyrLysGlyCysLeuIleArgGlnAspProSerGlyAlaPhe 434
 DB 2 ACTCTCTGGCCCGCCGACATA-NTNGCTGCTCATNCCGACAGACCCCGCGGCTTTC 60
 OY 435 SerLeuValGlyLeuSerGlnProHisArgSerLeuArgGlyLeuAlaIaAcysTrp 454
 DB 61 TCCCTGGTGGCTCCTAGTCAGCCACAGAACCTGCGGAGCTCTTGCACCTGCTGG 120
 OY 455 AsnSerGlyLeuArgValAspGlyAlaIleLeuTyrLeuThrSerCysAlaProArg 474
 DB 121 AATTCTGGGCTGCGAGTACAGCTGCTGCTGAACTACATCTGCTGCTCCAGAA 180
 OY 475 ProLysGlyLysSerAsnLeuIleValAlaArgArgGlyCysAsnProAlaProAlaPro 494
 DB 181 CCCAGAGAAAGTCCAAATTGATGCTGGTGGAGGGGCTGACCCCGCGCTGCCCT 240
 OY 495 GlyCysSerProSerCysAlaLeuThrGlnLeuSerPheHisThrIleProThrAsp 514
 DB 241 GCGTCTCCCGCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
 OY 515 SerLeuGlyLysGlyLysAsnLeuGlyLysSerPheHisThrIlePheArgGlySer 534
 DB 301 AGCCGAGAGTGGCAGAGAACCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
 OY 535 ArgArgGlyValAlaValAspGlyLysThrHisAspSerGlyValLeuLysValMetAsp 554
 DB 361 AGGCGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
 OY 555 SerArgHisArgAsnGlyMetGlySerPheLeuGlnAlaIleSerLeuMetSerGlnVal 574
 DB 421 TCCACACATCGAGACTGATGAGCTTTCTTGGAGACCCGCAAGCTTATGAGCCAACTA 480
 OY 575 SerTyrProHisLeuValLeuLeuHisGlyValCysMetAlaGlyAspSerIleMetVal 594
 DB 481 TCCTACCCGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
 OY 595 GlnGluPheValTyrLeu 600
 DB 541 CAGGAATGTGTATCTA 558

RESULT 13
 AJ442008 769 bp mRNA linear EST 19-APR-2002
 LOCUS AJ442008 dktf426 Gallus gallus cDNA clone 17j3r1, mRNA sequence.
 DEFINITION AJ442008
 ACCESSION AJ442008
 VERSION AJ442008.1 GI:20209229
 KEYWORDS EST.
 SOURCE chicken.
 ORGANISM Gallus gallus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 Phasianinae; Gallus.
 1 (bases 1 to 769)
 Buerstedde, J.M.
 TITLE Gallus gallus bursal lymphocyte EST
 JOURNAL Unpublished (2002)
 CONTACT Buerstedde JM
 COMMENT Cellular Immunology
 Heinrich-Pette-Institute
 Martinistr. 52, 20251 Hamburg, Germany
 Email: URL: http://genetics.hpi.uni-hamburg.de/dt40est.html.
 Location/Qualifiers
 1..769
 /organism="Gallus gallus"
 /strain="CB"

/db_xref="taxon:9031"
 /clone="17j3r1"
 /clone_lib="dktf426"
 /tissue_type="Bursa Of Fabricius"
 /cell_type="bursal lymphocyte"
 /dev_stage="2-3 weeks old"
 BASE COUNT 150 a 222 c 262 g 134 t 1 others
 ORIGIN

Alignment Scores:

Pred. No.: 2e-81 Length: 769
 Score: 906.00 Matches: 182
 Percent Similarity: 81.27% Conservative: 22
 Best Local Similarity: 72.51% Mismatches: 42
 Query Match: 15.46% Indels: 5
 DB: 9 Gaps: 3

US-09-397-967-16 (1-1099) x AJ442008 (1-769)

OY 770 ArgAspLeuAsnGlyLeuIleThrSerAspTyrGlyLeuLeuSerAspProThrProGly 789
 DB 1 CCGGACATCAACAGGCTCATCTCCGACATACAGCTCTCTCAGAGCTGCACCCGCG 60
 OY 790 IleProSerProArgAspGlyLeuCysValAlaGlyAlaGlnLeuTyrAlaCysGlnAsp 809
 DB 61 GATGTAGCGCTGCGG---GAGAGCTGCTGGGGGTACAGACAGCTGGCGGGGCGACGGC 117
 OY 810 ProAlaIlePheGlnGlnArgHisLeuLysTyrIleSerLeuLeuGlyLysGlnPhe 829
 DB 118 CCGGCTCAGTTCTGAGAGAGAGGACCTCAAGTACATCTCATGCTGGCAAGGCAACTTT 177
 OY 830 GlySerValGlyLeuCysArgTyrAspProLeuGlyAspAsnThrGlyProLeuValAla 849
 DB 178 GGGAGCGTGGAGAGCTTGGCGCTACAGACCCGCTGGGTGACAGCAGCGGTGAGCTGGTGC 237
 OY 850 ValLysGlnLeuGlnHisSerValProAspGlnGlnArgAspPheGlnArgGlnIleGln 869
 DB 238 GTGAAGAGCTGCACAGAGATTCGGCCAAAGGAGCTGACAGACTTTGAGAGGAGATCCAG 297
 OY 870 IleLeuLysAlaLeuHisSerAspPheIleValLysTyrArgGlyValSerTyrGlyPro 889
 DB 298 ATCTGCACTGCTGCTGCACAGAGCTTCACTGCAAGTACCGGGGCTGCTGCTACAGCCCT 357
 OY 890 GlyArgGlnSerLeuArgLeuValMetGlyTyrLeuProSerGlyCysLeuArgAspLeu 909
 DB 358 GGGCGCGGGGCTGCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 417
 OY 910 LeuGlnArg-----HisArgGlyLeuHisThrAspArgLeuLeuPheHisArgPheGln 927
 DB 418 CTGCAAAAGAACCAAGCAGCCG---CTGGAGCACCCGACGCTGCTGCTGCTGCTGCTGCTG 474
 OY 928 IleCysLysGlyMetGlyTyrLeuGlyAlaArgArgCysValHisArgAspLeuAlaIa 947
 DB 475 ATGTCCAAAGGCGATGAGTACTGCGGCGCGAGCCCTCGTGCACCGGAGCTTGCGCAGC 534
 OY 948 ArgAsnIleLeuValGlySerGlnAlaHisValLysIleAlaAspPheGlyLeuAlaLys 967
 DB 535 AGGAACATCTGCTGGAGAGAGGAGACCCAGTCGCAAGATCGGTGAGCTTGGGGTGGCAG 594
 OY 968 LeuLeuProLeuGlyLysAspTyrTyrValAlaArgGluProGlyGlnSerProIlePhe 987
 DB 595 CTGCTGCCAGAGACAGAGACTTACTGCTGGCGAGGAGCCCGGCGACAGCCCTCTTTC 654
 OY 988 TrpTyrAlaProGluSerLeuSerAsnIlePheSerArgGlnSerAspValIlePheSer 1007
 DB 655 TGGTACGACCCGAGTCCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 714
 OY 1008 PheGlyValValLeuTyrGlnLeuPheThr 1017
 DB 715 TTGGGGGTGCTGCTGCTGATGAGCTCTTCACC 745

RESULT 14
 BE916148

LOCUS BE916148 545 bp mRNA linear EST 29-SEP-2000
 DEFINITION 60166673F1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:3966903 5',
 mRNA sequence.
 ACCESSION BE916148
 VERSION BE916148.1 GI:10416498
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 545)
 NIH-MGC http://mgc.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Gilbert Smith, Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: L1AM9140 row: m column: 16
 High quality sequence stop: 545.
 Location/Qualifiers
 1..545
 /organism="Mus musculus"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone_image="IMAGE:3966903"
 /clone_lib="NCI_CGAP_Mam1"
 /tissue_type="tumor, biopsy sample"
 /dev_stage="10 months, virgin"
 /lab_host="DH10B"
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: Salt;
 Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt.
 Library constructed by Life Technologies. Investigator
 providing samples: Gilbert Smith, NIH"
 BASE COUNT 93 a 176 c 166 g 110 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 6.25e-79 Length: 545
 Score: 879.00 Matches: 175
 Percent Similarity: 97.25% Conservative: 2
 Best Local Similarity: 96.15% Mismatches: 3
 Query Match: 15.00% Indels: 3
 DB: 12 Gaps: 0
 US-09-397-967-16 (1-1099) x BE916148 (1-545)
 Oy 307 ProLaAGlGyIuHtSArGleuValThrValThrArgMetAspGlyHtSileuGluAla 326
 |||||||
 Db 3 CCGGAGGAGGAGCCCGGCTGCTACTGTCAACGATGACGCGCCACATCTGGAAGCG 62
 |||||||
 Oy 327 GluPheProGlyLeuProGluAlaLeuSerPheValAlaLeu-ValAspGlyTyrPheAr 346
 |||||||
 Db 63 GAGTTTCGGGGGCTGCTGAGGCGCTGTCTTGTGGCCCTCGGTGATGGTACTTCG 122
 |||||||
 Oy 346 GluLeuIleCysAspSerArgHtIstYrPheCysLysGluValAlaProProArgLeuLeu 366
 |||||||
 Db 123 CCGATCTCGGACTCCAGGCAATATTCTGCAAGAGAGTGGCGCGCCAGCGCTGTGA 182
 |||||||
 Oy 366 uGluGluAlaAspValCysHtS-GLYProIleThrLeuAspPheAlaIleHtSylLeuL 386
 |||||||
 Db 183 GGAAGAGGCGAGCTGTGACAGTGGACCATCAGTTAGACTTGGCCATCCACAAGCTGA 242
 |||||||
 Oy 386 ySaLaIaAGlYSerLeuProGlyTThrTyrIleLeuArgSerProGlnAspTyrAsp 406
 |||||||
 Db 243 AGGCGCTGGCTCCCTCCAGGACATATTTCTCCGCCGAGCCGACAGACTATGACA 302
 |||||||
 Oy 406 eRpHeuLeuThrAlaCysValGlnThrProLeuGlyProAspTyrIleGlyCysLeuI 426
 |||||||

|||||
 Db 303 GCTTCTCTTACCGCTGGCTCCAGACTCTCTTG6CCCCGACACAAAGGCTGCCTCA 362
 |||||||
 Oy 426 leArgGlnAspProSerGlyAlaPheSerLeuValGlyLeuSerGlnProHtSArSerL 446
 |||||||
 Db 363 TCCGCGAGACCCCGAGGGGGCTTCTCCCTGTGGCTCCACAGCCCAAGAGGCC 422
 |||||||
 Oy 446 euArgGluLeuAlaAlaCysTrpAsnSerGlyLeuArgValAspGlyAlaAlaLeuT 466
 |||||||
 Db 423 TCGGGAGCTGCTTGCAGCTCTGGAATTTCTGGCTGCGAGTGAAGAGTGTGCTGA 482
 |||||||
 Oy 466 ylrLeuThrSerCysAlaProArgProLysGluLysSerAsnLeuIleValAlaArgA 486
 |||||||
 Db 483 ACTAACATCTCTGCTGCTCCAG-CCCAAGGAAAAAGTCAATTGATGTGTGCGAA 541
 |||||||
 Oy 486 rg 486
 ||
 Db 542 GG 543
 ||
 RESULT 15
 BE913247 539 bp mRNA linear EST 29-SEP-2000
 LOCUS BE913247
 DEFINITION 601668246F1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:3968386 5',
 mRNA sequence.
 ACCESSION BE913247
 VERSION BE913247.1 GI:10410671
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 539)
 NIH-MGC http://mgc.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Gilbert Smith, Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: L1AM9144 row: k column: 11
 High quality sequence stop: 539.
 Location/Qualifiers
 1..539
 /organism="Mus musculus"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone_image="IMAGE:3968386"
 /clone_lib="NCI_CGAP_Mam1"
 /tissue_type="tumor, biopsy sample"
 /dev_stage="10 months, virgin"
 /lab_host="DH10B"
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: Salt;
 Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt.
 Library constructed by Life Technologies. Investigator
 providing samples: Gilbert Smith, NIH"
 BASE COUNT 92 a 160 c 149 g 138 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1.98e-78 Length: 539
 Score: 874.00 Matches: 174
 Percent Similarity: 97.21% Conservative: 0
 Best Local Similarity: 97.21% Mismatches: 5
 Query Match: 14.91% Indels: 2
 DB: 12 Gaps: 0
 US-09-397-967-16 (1-1099) x BE913247 (1-539)

```
OY 4 ProSerGIuIuThrProLeuIleProGIuArgSerCysSerLeuSerSerSerGIuAla 23
Db 3 CCAAGTAGAGAGACCTGTGATCCCTCAGAGCTTTCAGACCTCTCATCTCAGAGCA 62
OY 24 GlyAlaLeuHisValLeuLeuProProArgGlyProGlyProProGlnArgLeuSerPhe 43
Db 63 GGAGCCCTGCATGTCCTTCCTCCCGGGGAGCTGGGCTCCCGACGATAGTCATTTC 122
OY 44 SerPheGIuAspTyrLeuAlaGluAspLeuCysValArgAlaAlaLysAlaCysGlyIle 63
Db 123 TTC-TTTGGGACACTCTGCTGAGGATTTATGTGCCAGCTGCCAAGGCTGTGGCATC 181
OY 64 LeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSerCysTrpPhePro 83
Db 182 CTGCCGTATTATCATTCGCTTTTCGCTCTGGCCACTGAGGACTTCTCTGTGTTTCCC 241
OY 84 ProSerHisIlePheCysIleGluAspValAspPheGlnValLeuValTyrArgLeuArg 103
Db 242 CCAAGCCACATCTTCTGCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
OY 104 PheTyrPheProAspTrpPheGlyLeuGluThrCysHisArgPheGlyLeuArgLysAsp 123
Db 301 TTTTATTTCCCTGACTGTGTTGGGCTGAGACATGTCACCGCTCTGGGCTGCCAAGAT 360
OY 124 LeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLeuPheAlaGlnHisArgSer 143
Db 361 TTGACCAgTGCATCTTGCATACATGTTTAGAACATCTCTTGTGCTCAGCACCCAGAT 420
OY 144 AspLeuValSerGIuArgLeuProValGlyLeuSerMetLysGluGlnGlyIuPheLeu 163
Db 421 GACCTGGTAGTGCGGCGCTCCGCTGGGCTTAGCATGAGAGAGAGAGAGAGAGAGAG 480
OY 164 SerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGlnAlaGlnArgProGly 182
Db 481 AGCCTGGCCCTGCTGACTTGCAGATGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAG 537
```

Search completed: April 28, 2003, 16:55:17
Job time : 3456.75 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 28, 2003, 13:05:49 ; Search time 585.936 Seconds
(without alignments)
4223.916 Million cell updates/sec

Title: US-09-397-967-16

Perfect score: 5860

Sequence: 1 MAPSEPTPLIPORNSCLSS.....RPAFATLSPLDPLMRGRPG 1099

Scoring table:

BLOSUM62	Xgapop 10.0 , Ygapext 0.5
	Xgapop 10.0 , Ygapext 0.5
	Fgapop 6.0 , Fgapext 7.0
	Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL-frame+ p2n model -DEV-xlp
-Q=/cgnr_1/USPFO.spool/US09397967/rumat_26042003_182312_10702/app-query.fasta.1.1486
-DB=/geneseq_101002 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS-human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pro -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09397967.ecgn.1.1418 -rumat_26042003_182312_10702 -NCPU=6 -ICPU=3
-NO_XLPHY -NO_MMAR -LARGEDUJRY -NEG_SCORES=0 -MAIT -DSHLOCK=100 -LONGLOG
-DEV.TIMEOUT=10 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6
-Fgapext=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database:

- 1: Geneseq_101002.*
- 2: /SID2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
- 3: /SID2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
- 4: /SID2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
- 5: /SID2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
- 6: /SID2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
- 7: /SID2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
- 8: /SID2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
- 9: /SID2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
- 10: /SID2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
- 11: /SID2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
- 12: /SID2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
- 13: /SID2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
- 14: /SID2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
- 15: /SID2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
- 16: /SID2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
- 17: /SID2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
- 18: /SID2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
- 19: /SID2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
- 20: /SID2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
- 21: /SID2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
- 22: /SID2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
- 23: /SID2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
- 24: /SID2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5223	89.1	4016	AA11083	Mouse JAK3 protein
2	4418	75.4	3807	AA130862	Protein tyrosine k
3	2629	44.9	3435	AAV61801	JAK2 protein encod
4	2621	44.7	4482	AAAX80971	Human JAK2 kinase
5	2621	44.7	4482	AAZ58947	Human JAK2 kinase
6	2621	44.7	5117	AAAD24311	Human JAK2 kinase
7	2619.5	44.7	3629	AAQ85412	Human JAK2 kinase
8	2619.5	44.7	3629	AAAC6244	JAK2 polynucleotid
9	2619.5	44.7	3629	AAAC6245	Murine (Janus kina
10	2376.5	40.6	3473	AAO25307	JAK2 encoding DNA
11	2360	40.3	3495	AAAD2680	Murine JAK2 tyrosi
12	1908.5	32.6	4191	ABK72331	DNA encoding lymph
13	1904	32.5	3429	AAQ85413	Human JAK1 kinase
14	1904	32.5	3429	AAAC6245	Human JAK1 kinase
15	1904	32.5	3429	AAAD03608	Human (Janus kinas
16	1904	32.5	3540	AAO26039	JAK1 encoding DNA
17	1904	32.5	3541	ABK84065	Human CDNA differe
18	1904	32.5	3541	ABN95881	Gene #2379 used to
19	1904	32.5	3541	ABK72301	Lymphoma associate
20	1904	32.5	3541	ABL61826	Colon adenocarcino
21	1842	31.4	3538	ABK72332	DNA encoding lymph
22	1817.5	31.0	4176	ABK84572	Human CDNA differe
23	1817.5	31.0	4176	AAAD24312	Human TYK2 DNA. H
24	1815.5	31.0	3561	AAQ85414	Human TYK2 kinase
25	1815.5	31.0	3561	AAAC6246	Human TYK2 kinase
26	1815.5	31.0	3561	AAAD03609	Human polynucleoti
27	1813.5	30.9	4248	AAK52035	Human E2F-derived
28	1806.5	30.8	4218	AAH98351	Human polynucleoti
29	1806.5	30.8	4220	AAK53019	Human polynucleoti
30	1564	26.7	4294	AAAS8150	DNA encoding novel
31	1010	17.2	795	AAAS10808	Human Janus kinase
32	948.5	16.2	726	AAAS10807	Human Janus kinase
33	857.5	14.6	778	AAAS10809	Human Janus kinase
34	816.5	13.9	2007	AAK78092	Human immune/haema
35	746	12.7	2327	AAZ33632	Human breast tumou
36	713	12.2	2338	AAFL8071	Lung cancer associ
37	680.5	11.6	5045	ABLO2929	Drosophila melanog
38	630	10.8	1026	AAAS85149	DNA encoding novel
39	592.5	10.1	811	AAAS10806	Human Janus kinase
40	507	8.7	3222	AAK78095	Human immune/haema
41	497	8.5	3127	AAAS80650	DNA encoding novel
42	483.5	8.3	975	ABK72276	Lymphoma associate
43	483.5	8.3	975	ABK72333	DNA encoding lymph
44	482	8.2	3417	ABL91665	Human polynucleoti
45	477.5	8.1	3713	AAO47065	He truncated rece

ALIGNMENTS

RESULT 1
AA11083 standard; CDNA: 4016 BP.

08-APR-1996 (first entry)

Mouse JAK3 protein-tyrosine-kinase gene.

Mouse: Janus kinase; JAK3; protein-tyrosine-kinase; cytokine;
signal transduction; 3DPC13; reverse transcription; PCR;
polymerase chain reaction; primer; interleukin-3;
granulocyte-macrophage colony stimulating factor; CDNA probe;
granulocyte colony stimulating factor; Escherichia coli;
haematopoietic cell; differentiation; leukaemia; therapy; ss.
Mus musculus.


```

Db 1384 CGTGTGCGACCTAGCGCGCTAGCTGTCTCTTTCTTCCTAGCTTTTCCAGACCTTCTGT 1443
Qy 290 AAPPHERGGLIULLEVALASPVASERLLEASGLNALAPROARVALGIYPROALAGLY 309
Db 1444 GACTTTTCGGAAATCGGTGGATGTCACATCAACGACG--CCACGTGTGGGTCCGGAGGG 1501
Qy 310 GLNHISARGLEVALTHVALTHARGMETASPGLIHISLLEUGLUALIUPHEPPO 329
Db 1502 AG-CACCGCTGTACTGTGACACAGGATGACGCGCACATCTCGAAGCGGAGTTCCG 1560
Qy 330 GLYLEUPROGLUALLEUSERPHEVALALAEUVALASPGLYTHRPHARGLEULIECYS 349
Db 1561 GGGCTGCTGACGCGCTGTTCTTCTGCGCCCTGATGGAGTACTTCCGCGCTGATCTGC 1620
Qy 350 AAPPSEARHISHTYRPHESLYSGIYVALALAPROPARGLLEUGLUGLUALA 369
Db 1621 GACTCCAGGCAATTAATTTCTGCAAGAGGTGGCGGCGCACGCTGTGGAGGAGGCG 1680
Qy 370 ASPVALCYSHIAGLYPROILETHR----- 377
Db 1681 GAGCTGTGCATGACCCATCAAGTAAAGAGCTAGTTGGGTCAACATGGGGATGGGA 1740
Qy 378 -----LEUASPH 380
Db 1741 TGGGACGAGGACTGTCTGGGTGGACCTGCTGTCTTGGCCATGAGTTAGACTTT 1800
Qy 381 ALAILEHISLYSLEULYSALALAGLYSERLEUPROGLIYTHRTHLEULNARGRARGSER 400
Db 1801 GGCATCCACAGCTGAAGGCCCTGCTGCTCCAGGACCTATATTTCCGCGCGCAC 1860
Qy 401 PROGLNAPRTYRASPSPHELEULEUTHRALCYSAVALGINTHPROLEUGLYPROASP 420
Db 1861 CGGAGGACTATGACAGCTTTCTTACCGCTGCTCCAGACATCCCTGTGGCCCGGAC 1920
Qy 421 TYRYSGLCYALEULEILEARGINLEUPROSERGLIALAPHSERLEUVALILEUSER 440
Db 1921 TACAAAGGCTGCTCTATCCGCGAGGACCCAGCGGGGCTTCTCCCTGCTTGC-CTCAGC 1979
Qy 441 GLNPROHISARGSERLEUARGIULEULEUALALACYSYTRPASNSEGLIYLEUARGVAL 460
Db 1980 CAGCCCCACAGAAAGCTCGCGGAGGTGCTTGACG-TGCTGGAATTC-GGGCTGGGAGTA 2037
Qy 461 AAPPGLVALALALEUTHLEUTHRSERYCYSALAPROARGPROLYSGIYULYSERASN 480
Db 2038 GAGGGTGTGCTTAACTTAACATCTGTGCTCCAGACCCAGGAAAGAACAT 2097
Qy 481 LEULIIVALALARGAGLYCYASAPPROALAPROALAPROGLIYCYSESPROSECY 500
Db 2098 TTGATCGTGTGCGAAGGGGTGACACCCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2157
Qy 501 CYAALALEUTHINLEUSERPHEHISHTHRIEPTHASPSERLLEUGLUTRPHISGLU 520
Db 2158 TGTGGCTGACACACGCTGACCTTCCACAAATCCACAGGAGACGCTG---GCACACAG 2214
Qy 521 ASNLLEUGLIHISGLYSERPHETHRYSIIEPHEARGGLYSEARARGGLIYVALIASP 540
Db 2215 AACCTGGGTACGCTTCTTTTACCAAGATCTTCCGTGCGCGGAGGAGGAGTCTGTGAT 2274
Qy 541 GLYIUTHRHISASPSEGLIYVALILEULEULYSVALMETASPSERAIHISARGASNYS 560
Db 2275 GGTGAGACACATGACTCGGAAGTCTCTGAAGGTCACTGAGCTCAGACATCGGAACGTC 2334
Qy 561 METGLUSERPHELEUGLUALALASERLEUMETSERGINVALSETYRPHCHISLEUVAL 580
Db 2335 ATGGAGTCTTTCTGGAAGCGGCAAGCTTGTGAGACCAAGATCTTACCCGACCTGCTG 2394
Qy 581 LEULEUHLISGLIYVALCYMETALAGLYASPSERILEMETVALINGLUPHEVALIYTHLEU 600
Db 2395 TTACTGACGCGGCTGTGATGGCTGAGACACATCAATGAGTGTGAGAAATTTGTGTATCTA 2454
Qy 601 GLYALALEAPMETYRLEUARGLYSARGLIYHISLEUVALSERALASERTYRPLYSLEU 620

```

```

Db 2455 GAGACATTTGACATGTACTGCGCAAGCTGGCCACCTGGTGTACGCCACTGGAACGTG 2514
Qy 621 GLNVALTHRLYSGLINLEUALATYRALALEUANSTYRLEUGLIASPLYSGLIYLEUPROHIS 640
Db 2515 CAGGTGACCAAGAGCTGCGCATATGCTTTAACTACTTGGAGACCAAGCCCTTCTCTAC 2574
Qy 641 GLYASVALSERALARGLYSVALLEULEUALIARGGLIYGLIYASPGLYASPROPO 660
Db 2575 GGCACGTCCTCAGCAAGGAGGTGCTCTGCTGCTGAGAGGGGGGTGATGGAAATCCACT 2634
Qy 661 PHEILELYSEUSERASPROGLIYVALSERPROTHRALLEUSERLEUGLUNLEUTHR 680
Db 2635 TTCAATTAAGCTAGTATCTGTGTGACCTCCACGTGTGTGAGCTGGAAATGCTCAAC 2694
Qy 681 ASPARGIIEPRTYRVALALAPROGLIYCYSEUGLINGLUALAGLINTHRLEUCYSLEUGLI 700
Db 2695 GACAGAAATACCTGTGGGTGGCCCCCAATGTCTCCAGAGAGGCTCAGACACTGTCTGGAG 2754
Qy 701 ALAASPLYSITRPGLYPHEGLYALATHRTHRTTSPGLIYVALPHEGLNARGLYPROALHIS 720
Db 2755 GCTGACAGTGGGCTTTGGAGCCACACGTTGGAGGTGTTCACCGGGGACCCCCCAG 2814
Qy 721 ILETHRSELEUGLUPROALALYSLYSLEULYSPIETRYGLIYASPGLIYGLINLEUPRO 740
Db 2815 ATCACCCTGCTGAGCCCGCCAAAGCTGAAGTCTTATGAGACACAGGACAGCTGCTGCC 2874
Qy 741 ALALEULYITRPTHRLIYGLIYLEUALAGLYLEULEITHRGLIYCYMETALATYRASP 760
Db 2875 GCTTCAAATGACAGAACTGCGGGACTTATCACACAGTGCATGGCTATGATCTGCTGCC 2934
Qy 761 ARGARGPROSERPHEARGALALEULNARGSPLEUANGLIYLEUTHSERASP 780
Db 2935 CGGCGCCCTCTCTTCCAGAGTATCTTCCAGACCTCAACGGCTCATTAATCAATTAATCA 2994
Qy 781 GLULEULEUSERASPROTHRPROGLIYIIEPROSERPROARGASPGLYLEUCYSVALALA 800
Db 2995 GAGCTCTCTCCA-GACCCACACCTGTGATCCGAGTCTCTCGAATGATGACTGTGGCT 3049
Qy 801 GLYALAGLNLLEUTHYRALCYGGLNAPPROALALEPHEGLIYGLIYHISLEULYSTYR 820
Db 3050 GCGCGCCAGCTCTATGCTGTGCGAGGACCCCGCCATATC-GAGGAGAGACACTTAAGTAC 3108
Qy 821 ILESERLEULEUGLIYSGIYASNPHEGLYSEVALIYGLIYCYSEARGTYRASP 840
Db 3109 ATCTTTGTGGGCAAGGCACTTTGGCAGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3168
Qy 841 GLYSPASNTHRGLYPROLEUVALALAYALYSGLINLEUGLINHISSEVALPROASPGLN 860
Db 3169 ---GACAAATAGGGGACCCCTGTGTGCAATGAAACAGTACAGCACAGGCGGCGAGACCAG 3225
Qy 861 GLNARGASPPHEGLNARGGLIULIIEGLINLEULEULYSALALEUHLISSEARPHEILEVAL 880
Db 3226 CAGAGGAGCTTCCAGCGGGAGATTCAGATCTTAAAGCTCTGCAACAGGACTTATATGCTC 3285
Qy 881 LYSTYRARGGLIYVALSERTYRGLYPROGLIYARGLINSERLLEUARGLEUVALMETGLUTYR 900
Db 3286 AAGTACCGGGGAGTACAGTATGCGGCGCAGAGCTGCGGCTGGTGGTATGAGATGAC 3345
Qy 901 LEUPROSERGLIYCYSEULEARGASPLEULEUGLILNARGHISARG---GLYLEUHLISHTHASP 919
Db 3346 CTGCGCCAGCGCTCTGCGAGACTTCTTCCAGAGGCGCATGCGCGGCGCGCCAGACCGAG 3405
Qy 920 ARGLEULEULEUPHEALATRPGLINILECYSLYSGIYMETGLUTYRLEUGLIYALARGARG 939
Db 3406 CGCTTACTGTGCTTCCGTTGGCAGATGTCGAAGGCAATGGAGTACCTGGGTGCGCGCGCC 3465
Qy 940 CYSAVALHISARGASPLEUALALARGASNILEUVALIGLUSERGLIYALAHISVALYS 959
Db 3466 TGCCTACACGCTGACCTGCGCGCCGCAACATCTTCTGAGAGCGAGCTCATATGTAAG 3525
Qy 960 ILEALASPPHEGLYLEUALALATYSLEULEUPROLEUGLIYSGIYSPRYTYRVALIYVALARG 979
Db 3526 ATCGCGGACTTGGCGCTGCTAAGCTGCTGCGCTTGGGAAAGGACTTACTAGTGTGCTCGC 3585

```

QY 980 GluProGlyInSerProIlePheTrpTyrAlaProGluInSerLeuSerAspAsnIlePhe 999
 |||||
 Db 3586 GAGCTGGCCAAAGCCCATCTTTGGTATGCCCCGGAGTCCCATCTGACAAACATCTTC 3645
 |||||
 QY 1000 SerArgInSerAspValTTPSerPheGlyValValLeuTyrGluLeuPheThrTyrCys 1019
 |||||
 Db 3646 TCCCGCAATCTGACGTGGAGCTTGAGAGTGTGTGACGAGCTTCCACTACTGTC 3705
 |||||
 QY 1020 AsplSerCysSerProSerAlaGluPheLeuArgMetMetGlyProGluArgGluGly 1039
 |||||
 Db 3706 GACAGAGCTGCACCCATCGCTGAGTCTCTGCGCATGATGGGGCTGAGCCGTAAGGA 3765
 |||||
 QY 1040 ProProLeuGlySerArgLeuLeuGluLeuLeuAlaGluGlyArgGlyLeuProProPro 1059
 |||||
 Db 3766 CCCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3825
 |||||
 QY 1060 ThrCysProThrGluValGluGluLeuMetGluLeuGlySerTPAlaProGluProHisAsp 1079
 |||||
 Db 3826 ACCTGCCCCACCGAGTTTCAGAGAGCTCATGCACTGTGCGGGGCCAG-CCGCAAGAC 3884
 |||||
 QY 1080 ArgProAlaPheAlaThrLeuSerProGluInLeuAspProLeuTTPArgGlyArgPro 1098
 |||||
 Db 3885 CGGCAGGCTTCGCGACCCCTGAGCCGCCAGCTGAGCGGCTGTGGCGTAAGACCC 3941
 |||||
 RESULT 2
 AAT30862
 ID AAT30862 standard; DNA: 3807 BP.
 XX AAT30862;
 AC
 XX
 DT 13-SEP-1996 (first entry)
 XX
 DE Protein tyrosine kinase JAK3 CDNA.
 XX
 KW JAK3; protein tyrosine kinase; cell proliferation; differentiation;
 signal transduction; leukaemia; aplastic anaemia; myelodysplasia;
 polythemia vera; thrombocytosis; gene therapy; diagnosis; ss.
 KW
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 168..3359
 FT /*tag= a
 FT polyA_signal 3692..3696
 FT /*tag= b
 FT polyA_signal 3793..3797
 FT /*tag= c
 FT
 PN MO9618639-A1.
 XX
 PD 20-JUN-1996.
 XX
 PF 15-DEC-1995; 95WO-US16435.
 XX
 PR 15-DEC-1994; 94US-0357598.
 XX
 PA (UWJO) UNIV JOHNS HOPKINS SCHOOL MED.
 XX
 PI Clavin CI, Safford MG, Small D;
 DR WPI: 1996-300568/30.
 DR P-PSDB: AAR96037.
 XX
 PT Protein tyrosine kinase, JAK3, protein and nucleic acid - used in
 PT the gene therapy of cellular proliferative diseases, e.g. leukaemia,
 PT aplastic anaemia etc.
 XX
 PS Claim 4; Page 43-47; 97pp; English.
 CC A cDNA clone (AAT30862) codes for JAK3 (AAR96037), a new member of the
 CC JAK family of non-receptor protein tyrosine kinases, that probably
 CC plays a role in growth factor modulated differentiation.

CC proliferation and survival of haematopoietic stem/progenitor cells.
 CC It was derived from CD34+ mRNA obd. from normal human bone marrow
 CC by PCR amplification using primers (see also AAT30863-64) based
 CC on highly conserved motifs from protein tyrosine kinase catalytic
 CC domains. The JAK3 coding sequence can be incorporated into a
 CC vector and used for prodn. of recombinant JAK3. It can also be
 CC used in gene therapy protocols for leukaemia, myelodysplasia,
 CC polythemia vera, thrombocytosis and aplastic anaemia, or to
 CC stimulate haematopoietic cell proliferation. The gene was
 CC localised to chromosome 19, band p12-13.1.
 XX
 SQ Sequence 3807 BP; 716 A; 1219 C; 1099 G; 773 T; 0 other;
 Alignment Scores:
 Pred. No.: 0 Length: 3807
 Score: 4418.00 Matches: 870
 Percent Similarity: 86.948 Conservative: 62
 Best Local Similarity: 81.168 Mismatches: 130
 Query Match: 75.398 Indels: 12
 DB: 17 Gaps: 6
 US-09-397-967-16 (1-1099) x AAT30862 (1-3807)
 QY 1 MetaLabProSerGluGluThrProLeuIleProGluInArgSerCysLeuSerSer 20
 |||||
 Db 168 ATGCACTCCCAAGTGAAGAGAGCCCTGATCCCTCAGGCTTCAAGACCTTGTCTC 227
 |||||
 QY 21 SerGluInGluAlaLeuHisValLeuLeuProProArgGlyProGlyProProGluArg 40
 |||||
 Db 228 ACGAGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 287
 |||||
 QY 41 LeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuGlyValArgAlaAlaValAla 60
 |||||
 Db 288 CTATCTTCTCTCTTGGGAGCACTTGTGAGAGACCTGTGCGCTGTGCCAAGGCC 347
 |||||
 QY 61 CysGlyIleLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSerCys 80
 |||||
 Db 348 AGCGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 407
 |||||
 QY 81 TrpPheProSerHisIlePheCysIleGluAspValAspThr-GluValLeuValTyr 100
 |||||
 Db 408 TGGTT-CCCCGAGCCACATCTTCCGTGAGAGATCCAGACCCCAACATCTGTGTA 466
 |||||
 QY 100 ArgLeuArgPheTyrPheProAspTTPPheGlyLeuGluThrCysHisArgPheGlyLe 120
 |||||
 Db 467 CAGATTCGCTTTACTTCCCAATGTGGCTGAGAGAGCTTGGCAGCTTGGGCT 526
 |||||
 QY 120 uArgLysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLeuPheAlaG 140
 |||||
 Db 527 ACGCAAGATTTGGCCAGATGCTATCTTGAGCTCCAGTCTCGAGCACCTTTGCCA 586
 |||||
 QY 140 nHisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLysGluGlnI 160
 |||||
 Db 587 GCACGCGAGTGTGCTGAGTGGAGGCGCTCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 646
 |||||
 QY 160 yGluPheLeuSerLeuAlaValLeuAspLeuAlaGluMetAlaArgGluGlnAlaGlnAr 180
 |||||
 Db 647 TGAGTGTCTGACCTGCGCGCTGTGTGAGCTGCGCGGATGCGGAGAGACAGCGGCGG 706
 |||||
 QY 180 gProGlyGluLeuLeuTyrHisValSerTyrLysAlaCysLeuProProSerLeuArgAs 200
 |||||
 Db 707 GCGGGAGAGCTGCTGAGAGCTGCTGAGCTGCTGAGCTGCTGAGCTGCTGAGCTGCTGAGCT 766
 |||||
 QY 200 pValIleGlnGlyGlnAspPheValThrArgArgArgIleArgArgThrValValLeuAl 220
 |||||
 Db 767 CCGATTCACAGGCGCTGAGCTTGTGACGCGGAGGCGATTCGAGAGAGGCTGAGAGGCC 826
 |||||
 QY 220 aleuLeuProCysGlyArgLeuProGlyArgPro-TyrAlaLeuMetAlaLysTyrIle 240
 |||||
 Db 827 CCGTGC-CGGGGTGGCCCGCTGCGAGGACAGCGGACCTGCTCATGCGCAAGTACATCA 885
 |||||
 QY 240 euAspLeuGluArgLeuHisProAlaAlaThrThrGluThrPheArgValGlyLeuProG 260
 |||||

Db 886 TGGACCTGGAGCGCTGGATCCAGCCGGGGCCGAGACCTTCACGTGGGCTCCCG 945
Qy 260 IYALA-----GIngluGluProGlyLeuLeuArgValAlaGlyAspAsnGlyIleP 277
Db 946 GGGCCCTGGTGGCCACGACGGGCTGGGGCTGTCCGGCTGGTGGTGAACGGCGCATGG 1005
Qy 277 rGfPserSerAsnsp---GluLeuPheGlnThrPheCysAspPheProGluIleValA 296
Db 1006 CCTGGACCCAGGAGAACGAGAGTCCCTCCAGCCCTTCGCGACTTCACAAATGTAG 1065
Qy 296 sPValSerIleAsnGlnAlaProArgValGlyProIleGlyIleHisAlaGlyValThrV 316
Db 1066 ACATTAGCATCAAGCAGGCGCGCGGTGGCCGGCGGAGACCGCGCTGGTACTGG 1125
Qy 316 aLThrArmeLaspGlyHisIleLeuGluAlaGluPheProGlyLeuProGluAlaLeus 336
Db 1126 TTACGAGGACACAAACAGATTTTAGAGCGGAGTTCCAGGGCGTGGCCGAGGCTCT 1185
Qy 336 eRPhValAlaLeuValAspGlyTyrPheArgLeuIleCysAspSerArgHisTyrPheC 356
Db 1186 CATTGGTGGCGCTGGGACGGCTACTTCGGCTGCACGAGACTCCGACACTTCCTCT 1245
Qy 356 yALysGluValAlaProProArgLeuLeuGluGluAlaAspValCysHisGlyProI 376
Db 1246 GCAAGGAGGTG---GACCCGAGGCTGGTGGAGAGTGGCCGAGCATGCCACGGCCCA 1302
Qy 376 lEthrLeuAspPheAlaIleHisTysLeuValAlaAlaGlySerLeuProGlyIleThrI 396
Db 1303 TCACCTGTGACCTTTGGCATCAACAGCTCAGACGCTGGGGGCTCAGCTCGTGGCTCTAG 1362
Qy 396 lAeuAlaArgSerProGlnAspTyrAspSerPheLeuLeuThrAlaCysValGlnThrP 416
Db 1363 TTCTCGCGGCATCCCGCCAGGACTTGGACAGCTTCCTCCTCACTGCTGTGCCAAMAC 1422
Qy 416 rOleuGluProAspTyrIleGlyCysLeuIleArgGlnAspProSerGlyAlaPheSerL 436
Db 1423 CCTTGGCTCTATTATTAAGGCTGGCTCATCGGGCGACGCCACACAGAACCTTCCTTC 1482
Qy 436 euValGlyLeuSerGlnProHisArgSerLeuArgGluLeuLeuAlaAlaCysTyrPAsn 456
Db 1483 TGGTGGCTTCAGCCGACCCACAGACAGTCTTCGAGAGCTCCGGACCTGGGATGG 1542
Qy 456 eRGlYLeuAlaArgValAspGlyAlaAlaLeuTyrLeuThrSerCysAlaAlaProArgProL 476
Db 1543 GGGGCTGCACGTAGATGGGGTGGCAGTCACTTCCTGCTATGCCACAGCCA 1602
Qy 476 yAGluYLeuSerAsnLeuIleValAlaArgArgGlyCysAsnProAlaProAlaProGlyC 496
Db 1603 AAGAAAGTCCAACTGATTTGGTCCAGAGAGGTACAGCCACCCACATCTTCCTGG 1662
Qy 496 yASerPro---SerCysCysAlaLeuThrGlnLeuSerPheHisThrIleProThrAspS 515
Db 1663 TTCAGGCCCAATCCCAATACAGCTGAGTACAGATTTTCACAGATCCCGTGGTACA 1722
Qy 515 eRLeuGluThrPheIleGlnAsnLeuGlyHisGlySerPheThrIlePheArgGlySera 535
Db 1723 GCGTGAAGTGGATGAGAACTGGGGCATGGGCTCTTCACCAAGATTTACCGGGCTGTC 1782
Qy 535 rGArgGluValAlaAspGlyIleThrHisAspSerGluValIleLeuLeuValMetAspS 555
Db 1783 GGCATGAGGTGGTGGTGGGAGGGCCGAAGACAGAGTGGCTGGTGAAGGTATGGAG 1842
Qy 555 eRArgHisArgAsnCysMeGluSerPheLeuGluAlaAlaSerLeuMetSerGlnValS 575
Db 1843 CCAAGCAAGAACTGATGAGAGTCTTCCTGGAACACAGACAGCTTGATGAGCAAGTGT 1902
Qy 575 eRTyrProHisLeuValIleLeuHisGlyValCysMetAlaGlyAspSerIleMetValG 595
Db 1903 CCGTACCGGATCTGCTGCTGCTACAGGCGGTGCTGCTGAGAGACACACATGCTGCG 1962
Qy 595 lAGluPheValTyrLeuGluAlaIleAspMetTyrLeuArgTyrAspGlyHisLeuValS 615
Db 1963 AGGATTTGTACACCTGGGGGCGCATGATGATCTGCGAAACCTGGCCACCTGGTGC 2022

Qy 615 eRAlaSerTyrPheLeuGlnValThrTyrGlnLeuAlaTyrAlaLeuAsnTyrLeuGluA 635
Db 2023 CAGCCAGCTGGAGAGCTGACAGTGTGTCAAAACAGCTGGCCCTACGCCCTCACTATCTGAGG 2082
Qy 635 sPlySGlyLeuProHisIleGlyAsnValSerAlaArgLysValLeuLeuAlaArgGluGlyC 655
Db 2083 ACAAAAGCCCTGTCCCATGTGCAATGTCTGCGCGGAAGGTGCTCGGCTCGGAGAGGGG 2142
Qy 655 lYAspGlyAsnProProPheIleTysLeuSerAspProGlyValSerProThrValLeus 675
Db 2143 CTGATGGAGCCCGCCCTTCAATCAAGCTGATGATACCTGGGGTCAACCCCGCTGTATA 2202
Qy 675 eRLeuGluMetLeuThrAspArgIleProTyrValAlaProGluCysLeuGlnGluAlaG 695
Db 2203 GCGTGAGATGCTCACGACGAGATCCCTGGGTGGCCCGCGAGTGTCTCCGGAGAGCCG 2262
Qy 695 lInThrLeuCysLeuGluAlaAspLysTyrPheGlyAlaThrThrTyrGluValPheC 715
Db 2263 AGCACTTGAAGCTTGGAACTGACAAAGTGGGGCTTCGGCCGACGGTCTGGAGTGTATA 2322
Qy 715 lAArgGlyProAlaHisIleThrSerLeuGluProAlaLysLysLeuLysPheTyrGluA 735
Db 2323 GTGGCTCACCATGCCATCAAGTGGCCCTGATGCTCTGTAAAGAACTCCAAATTTATGAG 2382
Qy 735 sPGLnGlyLeuLeuProAlaLeuLysTyrPheGluLeuAlaGlyLeuIleThrGlnCysM 755
Db 2383 ACCGGCAGACAGCTGTGGCGCCCAAGTGGACAGAGCTGGCCCGTGGCATTCACACTGCA 2442
Qy 755 eRAlaTyrAspProGlyArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyL 775
Db 2443 TGGCCATATAGCGCGGTGACAGAGCCCTCTTACGAGCCCTTCATTCGTGACCTCAATAAGTC 2502
Qy 775 eLlThrSerAspTyrGlyLeuLeuSerAspProThrProGlyIleProSerProArgA 795
Db 2503 TCATCTCTCAGACTATGAGCTCTCTCAAGACACAC---TGGTCCCTGGGCATCTGTC 2559
Qy 795 sPGLuLeuCysValAlaGluAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGluG 815
Db 2560 ATGGGCTGTGG---AATGGGCCACACTCTATGCGGCCAAGACCCACAGATCTTGGAGG 2616
Qy 815 lAArgHisLeuLysTyrIleSerLeuLeuGlyLysGlyAsnPheGlySerValGluLeuC 835
Db 2617 AGAGACACTCAAGATCACTCACACCTGGGCAAGGGCTTTTGGCACCGTGGAACTGT 2676
Qy 835 yAArgTyrAspProLeuGlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGlnH 855
Db 2677 GCGGCTATGACCCGCTAGCGGACAAATACAGTGCCTGGTGGCCGTGAACAGCTCAGC 2736
Qy 855 lASerValProAspGlnGlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuH 875
Db 2737 ACAGCGGGCCAGACCGACAGAGGAGTTCAGGGGAGATTCAGATTCCTCAAAAGCACAG 2796
Qy 875 lASerAspPheIleValLysTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuA 895
Db 2797 ACAGTGAATTCATGTCAGATATCGTGTGCTGAGCTATGGCCCGGCCCAAGAGCCCTG 2856
Qy 895 rGLeuValMetGluTyrTyrLeuProSerGlyCysLeuArgAspLeuGlnArgHisArgG 915
Db 2857 CCGTGTCAATGAGTACCTGGCCAGCGGCTGTGGCGACTTCCTGCAAGCGGACCGGG 2916
Qy 915 lYLeuHisThrAspArgLeuLeuPheAlaTyrPGLnIleCysLysGlyMetGlnTyrL 935
Db 2917 GCGTGCATGCGACCGCCCTCTCTATTCCTCGGAGATCTGCAAGGGCATGAGTATACC 2976
Qy 935 euGlyAlaArgArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGluSerC 955
Db 2977 TGGGCTCCCGCGCTGCTGTCACCGGACCTGGCGGCCGCAAAACATCTCTGTAAGAGG 3036
Qy 955 lAAlaHisValLysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAspT 975
Db 3037 AGGCACAGCTCAAGATGCTGACTTGGCTTGAAGCTGCTGGCTTGAACAAAGACT 3096

QY 975 YTYrValValArgLupProGlyInSerProIlePheTrpTyrAlaProGlySerLeus 995
 |||||
 Db 3097 ACTACGTGGTCCGGAGAGCCAGAGCCCATTTTCTGGTATGCCCCCAATCCCTCT 3156
 QY 995 erAspaSnIlePheSerArgInSerAspValITrpSerPheGlyValValLeuTyrGluL 1015
 |||||
 Db 3157 CGGACACATCTCTCTCGCAGTCAGACGCTGTGAGCTTGGGGTGTCTCTGACGAGC 3216
 QY 1015 eupHeTrpTyrCysAspLysSerCysSerProSerAlaGluPheLeuArgMetLeuTyrP 1035
 |||||
 Db 3217 TCTTCACCTACTGTGGACAAAGGTGCAGCCCTGCGCGAGTTCCTGCGAGTGGAT 3276
 QY 1035 roGluArgLugLupProLeuGlyCysArgLeuLeuGluLeuAlaGluLupArgTyrL 1055
 |||||
 Db 3277 GTGAGCGGAGATGTCGCCGCTCTGCGCCCTCTTGGAACTCTGTGAGAGAGGCCAGAGCC 3336
 QY 1055 eupProProProProTyrCysProThrGluVal 1065
 |||||
 Db 3337 TGCGGGCGCTCTGCTGCTGCTGCTGCTGAGGTG 3368
 RESULT 3
 AAV61801
 ID AAV61801 standard; cDNA: 3435 BP.
 XX
 AC AAV61801:
 DT 20-JAN-1999 (first entry)
 XX
 DE JAK2 protein encoding cDNA.
 XX
 JAK2 protein: cytoplasmic domain; betac subunit; screening; asthma;
 KW Interleukin; granulocyte macrophage-colony stimulating factor; GM-CSF,
 KM IL-3; IL-5; human; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 10..3429
 FT /*tag= a
 FT /product= "JAK2 protein"
 XX
 PN WC9843087-A1.
 PD 01-OCT-1998.
 XX
 PE 23-MAR-1998: 98WC-US05387.
 XX
 PR 24-MAR-1997: 97US-0041511.
 XX
 PA (PHAA) PHARMACIA & UPJOHN CO.
 XX
 PI Kayles PS, Roberds SL;
 XX
 DR WPI: 1998-532151/45.
 DR P-PSDB: AAW76425.
 XX
 PT Screening for compounds useful for preventing or treating asthma -
 PT by determining if compounds inhibit binding of the JAK2 protein to
 PT e.g. IL-3, IL-5 or GM-CSF
 XX
 PS Disclosure: Fig 22A-C; 112pp; English.
 XX
 CC This cDNA encodes the full-length JAK2 protein. The N-terminal fragment
 CC of JAK2 protein can be used in the method of the invention of screening
 CC for compounds useful for treating or preventing asthma. The method
 CC comprises contacting a molecule comprising at least the N-terminal 294
 CC amino acid residues of the JAK2 protein, with another molecule comprising
 CC at least 13 membrane-proximal cytoplasmic amino acids of interleukin
 CC (IL)-3, IL-5 or granulocyte macrophage-colony stimulating factor (GM-CSF)
 CC proteins in the presence of the candidate compound, and determining
 CC whether the first and the second molecules form a complex. If the
 CC compound inhibits complex formation, it can be used to treat asthma.

SQ Sequence 3435 BP; 1157 A; 598 C; 733 G; 947 T; 0 other;
 Alignment Scores:
 Pred. No.: 4.05e-191 Length: 3435
 Score: 2629.00 Matches: 534
 Percent Similarity: 66.21% Conservative: 193
 Best Local Similarity: 48.63% Mismatches: 340
 Query Match: 44.86% Indels: 31
 DB: 19 Gaps: 14
 US-09-397-967-16 (1-1099) x AAV61801 (1-3435)
 QY 20 SerSerGluAlaGluValLeuLeuPro-ProArgGlyProGlyProGly 39
 |||||
 Db 111 TCCAGTCTTCAAGT-----GTATCTTACCATTCCCTTGGGAAATGTGAGGCGA 161
 QY 39 nArgLysSerPheSerPheGlyAspTyrLeuAlaGluAspLysCysValArgAlaAlaL 59
 |||||
 Db 162 TTATCTGACCTTTCACATCGGGAGTATGTGCAGAAATCTGTATCTGCTCTTA 221
 QY 59 sAlaCysGlyLeuLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPhe 79
 |||||
 Db 222 AGCTTGAGTATCACACCTGTGTATCATATATGTTCTTTATAGTGAACAAGAAAG 281
 QY 79 rCysTrpPheProProSerHisIlePheCysIleGluAspValAspThrGluValLeu 99
 |||||
 Db 282 GATCTGATATCCACCCACCATGCTTCCATATGATGATGATGATGATGATGATGATGAT 341
 QY 99 lTyrArgLeuArgPheTyrPheProAspTrpPhe-----GlyLeuGluThrCysHisAr 117
 |||||
 Db 342 CTACAGAAATAGATTATTTATCTTCTGTTATGTCAGTGCAGAGCAAGAGAGCTATCG 401
 QY 117 gPheGlyLeuArgLysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLe 137
 |||||
 Db 402 GCATGGAATATCTCGAGTGTGTAAGCTCCTCTTGTATGATCTTGTGATGATCTTACT 461
 QY 137 uPheAlaGluHisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetL 157
 |||||
 Db 462 CTTCCTTCAGTGGCGGATGATGTTTGTGATGATGATGATGATGATGATGATGATGATG 521
 QY 157 sGluGluGlyLupPheLeuSerLeuAlaValLeuAspLeuAlaGluLeuAlaArgLugL 177
 |||||
 Db 522 AACACAGGAGAAATGCTCTGGATGGAGTGTGATGATGATGATGATGATGATGATGATG 581
 QY 177 nAlaGluArgProGlyGlyLeuLeuLysThrValSerTyrLysAlaCysLeuProPhe 197
 |||||
 Db 582 CGATCAAAACCCACTGGCGCATCTATACCTATACAGCTACAGACATCTTACCAAAATG 641
 QY 197 rLeuArgAspValIleGluGlyGluAsnPheValThrArgArgIleArgArgThrVa 217
 |||||
 Db 642 TATTCAGCAAAAGATCCAGACTATCTATTTTCAAGGAAGCAATTAAGTACAGATT 701
 QY 217 lValLeuAlaLeuLeuProCysGlyArgLeuProGlyArgProTyrAlaLeuMetAla 237
 |||||
 Db 702 TCGCAGATTATTCAGCAATTCAGCCAAATGCAAGCCACTGCCAANAATTGAAACTTAA 761
 QY 237 sTyrIleLeuAspLeuGluArgLeuHisProAlaAlaThrThrGluThrPheArgValG 257
 |||||
 Db 762 GTATCTTATAATCTGGAACCTGTGAGTGTCTTACACAGAGAAATTTGAAGTAA 821
 QY 257 yLeuProGlyVala-----GluGluGluProGlyLeuLeuArgValAlaGlyAs 273
 |||||
 Db 822 AGAACCTGGAGAGTCCCTTCAGGTGAGAGATTTTTCACACCATTTTAATAACTGGAAA 881
 QY 273 pAsnGlyIleProTrpPheSer-----AsnAspGluLe 284
 |||||
 Db 882 CGGTGAGATTAGTGTCAAGAGGAAACATAAAGAAAGTACAGACTGACAGACAGAGA 941
 QY 284 uPheGluThrPheCysAspPheProGlyIleValAlaSerIleAsnGluAlaProAr 304
 |||||
 Db 942 TTACAGTTAATTCGATTTCTTAATATATATGATGATGATGATGATGATGATGATGATG 1001
 QY 304 gValGlyProAlaGlyLupHisArgLeuValThrValThrArgMetAspGlyHisIleLe 324

```

Db 1002 AGAGCGT---TCAAATGAAGCCGGTTGTACTATCCATAAGCAAGATGGTAAATCT 1058
Qy 324 uGUlualagluPheProGlyLeuProGUlualaleuSerPheValAlalaleuValaspGlyTy 344
Db 1059 GGAATTGCACTTACCTCATTAAGGAAAGCTTGTCTTCCGTCATTAAATTAATGATGATA 1118
Qy 344 rPheArgLeuIleCysAspSerArgHisIleTyPheCysLysGUlualalaproProArgLe 364
Db 1119 TTATGATTAATACGACATGACATCATTTACCTCTGTAAAGAGATGACACCTCCAGCGCT 1178
Qy 364 uLeuGUlucGUlualaleuValCysHisGlyProIleThrLeuAspPheAlaIleHisIly 384
Db 1179 GCTTGAAATATACAAACCACTGATGCGCCCAATTCGATGAGATTTTGGCATAGTAA 1238
Qy 384 sLeuValAlalagIysLeuProGUlYThrTyrIleLeuArgAspSerProGUlaspGly 404
Db 1239 ACTGAGAAAGACGGATATACAGACTGACCTGATGTACTGATCCAGTCCTAAGGAGCTT 1298
Qy 404 rAspSerPheLeuLeuThrAlaCysValGlnThrProLeuGUlProAspTyrLysGlyCy 424
Db 1299 TAATTAATATATTTTGTGACTTTTTCGTGCGAGAGAAATGTCATTGAATATAAACACG 1358
Qy 424 sLeuIleArgGlnAspProSerGlyAlaPheSerLeuValGlyLeuSerGlnProHisAr 444
Db 1359 TTTGATTAACAAAAATGAGATGAGAGATACAACTCAGTGGACAAAGAAACACTGAG 1418
Qy 444 gSerLeuArgGUlLeuLeuAlalaleuCystrProAspSerGlyLeuArgValaspGlyAla 464
Db 1419 CAGCTTTAAAGATCTTTTGAATGTTTACCAGATGGAACCTGCTCCAGACAAATATAT 1478
Qy 464 aLeuTyrIleuThrSerCysAlaAlaProArgProLysGUlYsSerAsnLeuIleVala 484
Db 1479 TTTCCAGTTTACTAAATGCTGCCCCCAAGCCAAAGATTAATCAAACTTCTGTGCTT 1538
Qy 484 lArg---ArgGlyCysAsnProAlaProAlaProGUlYsSerProSerCysCysAlale 503
Db 1539 CAGAAGCAATGGTCTTGTGATGTACCAACTCACCAACATTACAGAGCCCTACTCATAT 1598
Qy 503 uThrGlnLeuSerPheHisIleProThrAspSerLeuGUlTyrPheHisGlnAsnLeu 523
Db 1599 GAAACCAATGGTGTTCACAAATGAGAAATGAGATTTGATTAATTAAGAAAGCCCTGG 1658
Qy 523 yHisGlySerPheThrIlePheArgGlySerArgArgGUlValasp---GlyG 542
Db 1659 CCAAGGCACTTTACAAAGATTTTAAAGCGCTACGAGAGAGATGAGAGACTAGGCTCA 1718
Qy 542 uThrHisAspSerGUlValLeuLeuLysValMetAspSerArgHisArgAsnCysMetG 562
Db 1719 ACTGCATGAACAGAAAGTTCTTTTAAAGTTCTGATTAAGCACACAGAACTATTTCAGA 1778
Qy 562 uSerPheLeuGUlAlalaleuSerLeuMetSerGlnValSerTyrProHisLeuValLeu 582
Db 1779 GCTTTTCTTTAAGCAGACAGATGATGATGAGCAAGCTTTCTCCAGACATTTGGTTTAA 1838
Qy 582 uHisGlyValCysMetAlagIYAsp---SerIleMetValGlnGUlPheValTyrLeuG 601
Db 1839 TTATGATATATCTGCTGCGAGAGACAGATATTTCTGCTGAGAGCTTTGTAAATTTGCG 1898
Qy 601 yAlaIleAspMetTyrLeuArgLysArgGlyHisLeuValSerAlaSerTrpLysLeuG 621
Db 1899 ATCTACATGATACATATCTGAAAAAGAAATTAATTAATTAATTAATGGAACCTTGA 1958
Qy 621 nValThrLysGlnLeuAlaIleValLeuAsnTyrLeuGUlYsAspLysGUlLeuProHisG 641
Db 1959 AGTGTCTAAGCAAGTTGGCATGGCCATCATTTTCTAGAAGAAACACCCCTTATTCATGG 2018
Qy 641 yAsnValSerAlaArgLysValLeuLeuAlaArgGlnGUlYs-----AspGlyAsnPr 659
Db 2019 GAATGTATGTGCAAAATATTTCTGCTTATCAAGACAGACAGACAGACAGAAATCC 2078
Qy 659 oProPheIleLysLeuSerAspProGUlYValSerProThrValLeuSerLeuGluMetLe 679

```

```

Db 2079 TCCTTTATCAAACTTAGATCTGGCATTAAGTATTAACAGTTTTCGCAAAAGCATTTCT 2138
Qy 679 uThrAspArgIleProTrpValAlaProGUlYsLeuGlnGUlAlaGlnThrLeuCysLe 699
Db 2139 TCAGAGAGAAATACCATGGGTATCCACCTGAAATGATGCAATTAATTAATTAATTT 2198
Qy 699 uGUlualaspLysTrpGlyPheGUlYAlaIleThrThrTrpGUlValPheGlnArgIleProAl 719
Db 2199 GGCACAGAGAAATGAGATTTGGTACACACTTGTGGGAAATCTGCAGTGGAGAGATTA 2258
Qy 719 aHisIleThrSerLeuGUlProAlaLysLeuLysLeuAspPheTyrGUlYsPheGlnGly 739
Db 2259 ACCCTAAGTGTCTGATTTCTCAAAAGAAAGCTTCAATTTTGAAGATGACATCAGCT 2318
Qy 739 uProAlaLeuLysTrpTrpHisLeuAlagIYLeuIleThrGlnCysMetAlaTyrAspTr 759
Db 2319 TCCTGCACCAAAAGTGGCAGAAATTAACAACTTTTAATTAATTTGATGATTAAGAAC 2378
Qy 759 oGUlYArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGUlYLeuIleThrSerAs 779
Db 2379 AGATTTCAAGCCCTTCTTCAGAGCCATGATACAGATGCTTAACAGTTTGTTAACGCA 2438
Qy 779 rTyrGUlLeuLeuSerAspProThrTrpGUlYLeuProSerProArg---AspGlyLeuCy 798
Db 2439 TTATGACATTAATTAACAGAA---AATGACATGTACCAAAATAGAGATGAGTGCCTAG 2495
Qy 798 sValAlagIYAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGlnGUlYArgHis 818
Db 2495 GTTTTCTGGTGGCC-----TTTGAAAGACCGGATCTCTACACAGTTTGAAGAGACATTT 2549
Qy 818 uLysTyrIleSerLeuLeuGUlYsGlyAsnPheGlySerValGlnLeuCysArgTyrAs 838
Db 2550 GAAATTTCTACAGCAACTTGGCAGAGGATTAATTTGGAGATGAGAGATGTCGCGGATGA 2609
Qy 838 rProLeuGUlYAspAsnThrGlyProLeuValAlaValLysGlnLeuGlnHisSerValPr 858
Db 2610 CCGCTTACAGGACCAACACTGGGAGGTGCTCCCTGTAAAGCTTTCACCAATAGCTGA 2669
Qy 858 oAspGlnArgAspPheGlnArgGUlYLeuGlnIleLeuLysAlaLeuHisSerAspRh 878
Db 2670 AGAGCACCTACAGACATTTGAAGGGAATTAATTCCTGAATTCCTACACATGACAA 2729
Qy 878 eIleValLysTyrArgGUlYValSerTyrGUlYProGUlYArgGlnSerLeuArgLeuValMe 898
Db 2730 CATGTGAAGATCAAGGAGGTGCTGACAGTCTGTCGGCGGATTAATTAATTAAT 2789
Qy 898 tGUlYTrLeuProSerGlyCysLeuArgAspLeuGlnArgHisArg---GlyLeuHis 917
Db 2790 GGAATATTTTCCATATGGAAGTTTACGAGACTATCTTCAAAAACATTAAGAACGGATGA 2849
Qy 917 sThrAspArgLeuLeuLeuPheAlaIleProGlnIleCysLysGlyMetGlyTrpLeuGUlY 937
Db 2850 TCACATTAACACTTCTGACGTACACATCTCAGATATGCAAGGGATATGAGATATCTTGATC 2909
Qy 937 rArgArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGlnSerGlnAlaHis 957
Db 2910 AAAAGATATATTCACAGGAGATCTGCAACAGAAATATATGTGTGGAGACAGACAGACAG 2969
Qy 957 sValLysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGUlYLysAspTyrTrpY 977
Db 2970 AGTTAAATTTGAGATTTTGGGTAAACCAACTCTTCCACAAAGCAAAAGATTAATTA 3029
Qy 977 lValArgGUlYProGUlYGlnSerProIlePheTrpTyrAlaProGUlYSerLeuSerAspAs 997
Db 3030 ACTTAAGAGACCTGGTGAAGTCCATATCTGATGCTGCAGAAATCTACACAGAGAG 3089
Qy 997 nIlePheSerArgGlnSerAspValTrpSerPheGlyValValLeuTyrGUlYLeuPheThr 1017
Db 3090 CAAGTTTCTGTGGGCTCAGATGTTTGGAGCTTTGGAGGCTTCTGTATGAACCTTTTCC 3149
Qy 1017 rTyrCysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetGlyProGUlYAr 1037
Db 3150 ATACATTTGAGAGAGTAAAGTCCACAGCGGAATTTATGCTATGATGATGCAATGACAA 3209

```

OY 1037 gGluGlyProProLeuLeu---ArgLeuLeuGluLeuLeuLaGluGlyArgArgLeuPr 1056
 Db 3210 ACAAGGACAGATGATCGTTCATTTGATAGAACTTTTGAGAAATATGAGCAATTAAC 3269
 OY 1056 opProProProthCysPProthrgLvalGlnGluLeuMetGlnLeuCysTrpAlaProG1 1076
 Db 3270 AAGACGAGATGATGCCAGATGATCTATATGATCATGACAGATGCTGGAACAAATA 3329
 OY 1076 uProHisAspArgProAlaPheAlaThrLeuSerProGlnLeuAspProLeu 1093
 Db 3330 TGTAAATCAACGCCCTTCCTTAGGATCTAGCTTCGAGTGCATCAATA 3381
 RESULT 4
 ID AAX80971 standard; cDNA; 4482 BP.
 XX AAX80971;
 DT 03-SEP-1999 (first entry)
 XX Human JAK2 kinase encoding cDNA.
 DE JAK2 kinase; arteriosclerosis; asthma; bronchitis; emphysema; psoriasis;
 KM inflammatory bowel disease; inflammation; osteoarthritis; oncogenesis;
 KW rheumatoid arthritis; septic shock; systemic lupus erythematosus;
 KM leukaemia; human; ss.
 XX Homo sapiens.
 OS US914393-A.
 PN 22-JUN-1999.
 PD 05-DEC-1995; 9505-0567508.
 PF 05-DEC-1995; 9505-0567508.
 PR 05-DEC-1995; 9505-0567508.
 XX (INCY-) INCYTE PHARM INC.
 PA Coleman R, Stuart SG;
 PI WPI: 1999-384188/32.
 DR P-PSDB: AAY21698.
 XX Polypeptides and amino acids useful for modulating human jak2 kinase activity
 PS Claim 2; Columns 23-28; 37pp; English.
 XX This cDNA encodes a human JAK2 kinase polypeptide. Host cells transformed
 CC with recombinant jak2 kinase nucleic acid are used for the recombinant
 CC production of the protein. Purified JAK2 may be used to produce
 CC antibodies or identify antagonists or inhibitors of JAK2. JAK2, anti-JAK2
 CC antibodies and JAK2 antagonists or inhibitors may be used to treat,
 CC prevent or diagnose conditions associated with altered or uncontrolled
 CC Jak2 expression, e.g. arteriosclerosis, asthma, bronchitis, emphysema,
 CC inflammatory bowel disease, inflammation, leukemia, oncogenesis,
 CC osteoarthritis, psoriasis, rheumatoid arthritis, septic shock and
 CC systemic lupus erythematosus.
 SO Sequence 4482 BP; 1477 A; 795 C; 953 G; 1249 T; 8 other:
 Alignment Scores:
 Pred. No.: 2,37e-190 Length: 4482
 Score: 2621.00 Matches: 532
 Percent Similarity: 66.03% Conservative: 193
 Best Local Similarity: 48.45% Mismatches: 342
 Query Match: 44.73% Indels: 31
 DB: 20 Gaps: 14
 US-09-397-967-16 (1-1099) x AAX80971 (1-4482)

OY 20 SerSerGluAlaGlyAlaLeuHisValLeuLeuPro-ProArgGlyProGlyProProG1 39
 Db 447 TCAGTCTTCTAGGT-----GTATCTTTACCATCTCCCTTGGAATCTGAGGAGA 497
 OY 39 nArgLeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysValArgAlaAla 59
 Db 498 TTATCTGACCTTCCATCTGGGAGATGTTGAGAAATCTGTATTCCTGCTCTAA 557
 OY 59 sAlaCysGlyLeuLeuProValTyrHisSerLeuAlaLeuAlaThrGluAspPhe 79
 Db 558 AGCTTGAGTGCACACCTGTGATCATATATATGTTGCTTTATATGATGACAGAAAG 617
 OY 79 rCysTrpPheProPheSerHisIlePheCysIleGluAspValAspThrGlnValLeu 99
 Db 618 GATCTGTATCTGACCCCAACCTGCTTCCATATGATGATGACCAACGCACTATGACT 677
 OY 99 lTyrArgLeuArgPheTyrPheProAspTrpPhe-----GlyLeuGluThrCysHisAr 117
 Db 678 CTACAGATTAAGATTTACTTCTCTGTTGATTCAGTGGCAGCAACAGAGCCCTATCG 737
 OY 117 gPheGlyLeuArgGlyAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisIle 137
 Db 738 GCATGGAATATCTGAGGTGCTGAAGCTCCTCTTGATGACTTTGTCATGCTTACT 797
 OY 137 uPheAlaGlnHisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMet 157
 Db 798 CTTTGCTCAGCGCGCATGATTTTGTCATGATGATGATGATGATGATGATGATGATG 857
 OY 157 sGluGlnGlyGluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGlu 177
 Db 858 AACACAGAAAGATGCTTGAGATGACAGTGTATGATATGATGATGATGATGATGATG 917
 OY 177 nAlaGlnArgProGlyGluLeuLeuLysThrValSerTyrLysAlaCysLeuProPhe 197
 Db 918 CGATCAAAACCCACTGGCGCATCTATACCTATACAGTACAGTACAGTACAGTACAGT 977
 OY 197 rLeuArgAspValIleGlnGlyGlnAspPheValThrArgArgGlyIleArgGlyThr 217
 Db 978 TATTCGAGCAAGATGCCAGTATCTTGTTCACAAAGCAAGCAAGTAAAGTACAGATT 1037
 OY 217 lValLeuAlaLeuLeuProCysGlyArgLeuProGlyArgProTyrAlaLeuMetAla 237
 Db 1038 TCGCAGATTTTTCAGCAATTCAGCAATTCAGCAATTCAGCAATTCAGCAATTCAG 1097
 OY 237 sTyrIleLeuAspLeuGluArgLeuHisProAlaAlaThrThrGluThrPheArgVal 257
 Db 1098 GTATCTTATAATCTCGAACTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1157
 OY 257 yLeuProGlyAla-----GlnGluGluProGlyLeuLeuArgValAlaGlyAs 273
 Db 1158 AGAACCTGGAAGTGCTCCTTCAGTGAAGAGATTTTTCGAACCATTTATTAATGACTGA 1217
 OY 273 pAsnGlyIleProTyrPheSer-----AsnAspGluLe 284
 Db 1218 CGGAGGAATTCAGTGTCACAGGAAACATAAGAAAGTACAGACATGACAGAAAGCA 1277
 OY 284 uPheGlnThrPheCysAspPheProGluIleValAspValSerIleAsnGlnAlaProAr 304
 Db 1278 TTACAGTATATGCGATTTCTCTAATATATGATGATGATGATGATGATGATGATG 1337
 OY 304 gValGlyProAlaGlyGlnHisArgLeuValThrValThrArgMetAspGlyHisIleLe 324
 Db 1338 AGAGGGT--TCAATGAAAGCGAGTTGATCATATCATATGATGATGATGATGATG 1394
 OY 324 uGluAlaGluPheProGlyLeuProGluAlaLeuSerPheValAlaLeuValAspGlyTy 344
 Db 1395 GGAATTCGAATTCAGTCAATTAAGGAAAGCTTTGCTTTCTGTCATTAATGATGATA 1454
 OY 344 rPheArgLeuIleCysAspSerArgHisTyrPheCysLysGluValAlaProProArgLe 364
 Db 1455 TTATAGATTAATCTGACATGATCATATACCTCTGTAAAGATAGACCTCCAGCGCT 1514
 OY 364 uLeuGluGluAlaAspValCysHisGlyProIleThrLeuAspPheAlaIleHisLys 384

RESULT 5

AAZ58947 standard; cDNA; 4482 BP.

AAZ58947;

03-MAY-2000 (first entry)

Human JAK2 kinase (HJAK2) encoding cDNA.

Janus family nonreceptor protein-tyrosine kinase-2; JAK2; HJAK2; human; signal transduction; arteriosclerosis; asthma; bronchitis; emphysema; inflammatory bowel disease; leukemia; oncogenesis; osteoarthritis; psoriasis; rheumatoid arthritis; systemic lupus erythematosus; cytostatic; osteopathic; dermatological; antibacterial; septic shock; immunosuppressive; ss.

Homo sapiens.

US6019966-A.

01-FEB-2000.

19-NOV-1998; 9805-0196480.

05-DEC-1995; 9505-0567508.

(INCY-) INCYTE PHARM INC.

Stuart SG, Coleman R;

WPI: 2000-146859/13.

P-PSDB; AAT77552.

Human Janus family nonreceptor protein-tyrosine kinases useful as diagnostic reagents and for preventing, diagnosing and treating diseases such as arteriosclerosis, asthma and leukemia.

Example 5; Fig 1A-F; 33pp; English.

This cDNA encodes a human Janus family nonreceptor protein-tyrosine kinase-2 (JAK2) polypeptide (HJAK2). The JAK2 polypeptides may be used as diagnostic reagents as they react with a range of target proteins including growth hormone, prolactin, erythropoietin and cytokine receptors. They may also be used for the production of antibodies specific for JAK2, which may be used to inhibit its activity and prevent or treat disorders associated with over expression of JAK2. Conversely, the JAK2 polypeptide may be administered to supplement the patients own production and counter mutations that may lead to the expression of an inactive enzyme. The protein may also be used to screen candidate reagents for modulators of JAK2 function. The antagonists and antibodies bind to the JAK2 protein and prevent the transfer of high energy phosphate molecules, therefore blocking signal transduction. Disorders that may be treated by administration of JAK2 polypeptides, anti-JAK2 antibodies and the agonists and/or antagonists. Include arteriosclerosis, asthma, bronchitis, emphysema, inflammatory bowel disease, leukemia, oncogenesis, osteoarthritis, psoriasis, rheumatoid arthritis, septic shock and systemic lupus erythematosus.

Sequence 4482 BP; 1477 A; 795 C; 953 G; 1249 T; 8 other;

Alignment Scores:

Pred. No.:	2,37e-190	Length:	4482
Score:	2621.00	Matches:	532
Percent Similarity:	66.03%	Conservative:	193
Best Local Similarity:	48.45%	Mismatches:	342
Query Match:	44.73%	Indels:	31
DB:	21	Gaps:	14

US-09-397-967-16 (1-1099) x AAZ58947 (1-4482)

QY 20 SerSerGluAlaGlyAlaLeuHisValLeuLeuPro-ProArgGlyProGlyProProG1 39

Db	447	TCACAGTTCTTCAGT-----GATCTTACCATTCCTGGGAATCTGAGCAGA	497
QY	39	naArgLeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuGlyValArgAlaAla	59
Db	498	TTATCTGACCTTCCATCTGGGAGATGTTGTGACAGAAATCTGTATCTGCTTCA	557
QY	59	saIacGlyGlyIleLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPhe	79
Db	558	AGCTTGTGATACACCTGTGTATCATATATGTTTGTATATGAGGAACAGAAAG	617
QY	79	rcYstrPheProPheSerHisIlePheGlyIleGluAspValAspThrGlnValLeu	99
Db	618	GATCTGTATCCACCCACACATGCTTCCATATAGTAGTAGTACACAGGACATATG	677
QY	99	lYrArgLeuArgPheTyrPheProAspThrPhe-----GlyLeuGluThrCysHis	117
Db	678	CTACAGAAATAGATTTTACTTCTGCTTGTTGATTCAGTAGTCCAGACACAGCC	737
QY	117	gPheGlyLeuArgGlyAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHis	137
Db	738	GCATGGAAATATCTCGAGGTGCTGAGCTCTCTTGTATGACTTGTATCTTACCT	797
QY	137	urPheAlaGlnHisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSer	157
Db	798	CTTTGCTCAGTGGCGGATGATTTTGTGATGATGATGATGATGATGATGATG	857
QY	157	sgLingLingLingLingLingLingLingLingLingLingLingLingLingLing	177
Db	858	AACACAGAGAAATGCTTGTGATGACAGTGTATGATGATGATGATGATGATG	917
QY	177	naIaGlnArgProGlyGluLeuLeuYstrValIserTyrLysAlaCysLeuPro	197
Db	918	CGATCAACCCACTGCGCATCTATACCTATACCTATACCTATACCTATACCT	977
QY	197	rIeArgAspValIleGlnGlyGlnAspPheValThrArgArgGlyLeuArgThr	217
Db	978	TATTCGACAAAGATTCACAACTATCTATTTGACAGAGAACGATAGTACAG	1037
QY	217	IvalLeuAlaLeuLeuProGlyGlyArgLeuProGlyArgProTyrAlaLeuMet	237
Db	1038	TCCGAGATTTATTCAGCAATTCAGCAACCAAGCCACGACCAAGCTTGA	1097
QY	237	sTyrIleLeuAspLeuLuarGlyLeuHisProAlaIleThrGluThrPheArg	257
Db	1098	GATCTTATATATCTGGAACCTGCACTGCTCTTACACAGAAATTTGACAT	1157
QY	257	YLeuProGlyAla-----GlnGluGluProGlyLeuLeuArgValAlaGly	273
Db	1158	AGAACCTGGAAGTGTCTTCAGGTGAGAGATTTTGCACCATTTATATAC	1217
QY	273	paSngLylIleProTyrPheSer-----AsnSglIle	284
Db	1218	CGTGGAAATTCAGTGCAGAGGAGGAACATTAAGAAAGTAGACACTGACAG	1277
QY	284	urPheGlnThrPheCysAspPheProGluIleValAspValSerIleAsnGln	304
Db	1278	TTTACAGTATATATGCGATTTCTTATATATGATGATGATGATGATGATG	1337
QY	304	gValGlyProAlaGlyGluHisArgLeuValThrValThrArgMetAspGlyHis	324
Db	1338	AGAGGT---TCAAATGAACCCAGATTGTATCATTCATTAAGCAAGATTA	1394
QY	324	uGluAlaGluPheProGlyLeuProGluAlaLeuSerPheValAlaLeuValAsp	344
Db	1395	GGAATTCGAATTCATTCATTAAGGAGGCTTCTCTTCTGATTAATGATGATA	1454
QY	344	rPheArgLeuLeuLeuAspSerArgHisTyrPheCysLysGluValAlaProArg	364
Db	1455	TTATGATTAATCAGTACATGACATCATTTACCTGTGTAACAAATAGACCC	1514
QY	364	uLeuGluGluGluAlaAspValCysHisGlyProIleThrLeuAspPheAlaHis	384

D	1515	GCCTGAAGAAATATACAAACCAACTGTCACTGGCCCAATTTCCATGGATTTTGGCATTAAGTAA	157
OY	384	stleuysalalaclyserleuproglythrtylleuaurargserproglinspy	404
D	1575	ACTGGAAGAAAGCAAGTAATCAGACTGAGGACCTGTATGACTTCGATTCGACTTAAGACACTT	1634
OY	404	rasperpheuleuethrhalacysvalglnthrproleuglyproasptrylsglycys	424
D	1635	TAATTAATATTTTTTGTGACTTTTGGCTGTGAGCGAGAAATATGTCATTGAATATTAACACAG	1694
OY	424	stleuilearglnasprprosergilyalapheserleuvalglyleuenserprohlsnr	444
D	1695	TTTGTGATTCACAAAATAATGAGATGAAGATCAACCTCAGTGGGACAAAGAAGACTTCAG	1755
OY	444	gserleuarglnleuileuallalacysttrpranserglyleuaurvalaspylvalal	464
D	1755	CAGTCTTAAAGATCTTTTGAATTTGTTACAGATAGCAAACTGTTCCGTGACGAACATTAAT	1815
OY	464	aleuutytleuethsercyscysalaproactprolysglyseraspheleuilevala	484
D	1815	TTTTCCAGTTTACTTAATGTGCTGTCCCCCAAGCCAAAAGATTAATCAACCTTCGTGCTT	1875
OY	484	lary---argcylcysasnrproalarproglycysserprosercyscysalale	503
D	1875	CAGAACGAAATGGTCTTCTGTGATACCAACCTCACCACACTTACAGAGGCTTAATCATAT	1934
OY	503	uthrclineuenserphenisthrileprothrasperleuileuthrphisglinsleucl	523
D	1935	GAAACCAATGGCTTTTACAAATCAGAAATGAAGATTTGATATTTAAAGAAAGCCCTTGG	1994
OY	523	yhslglyserphenethrlysllehearglyserafargduvalvalasp---glycl	542
D	1995	CCAAAGCAGCTTTTACAAAGATTTTAAAGCGCTACCAAGAGAAGTAGAGACTACAGCA	2054
OY	542	uthrhisasperserduvalleuileuysvalmetasperserduhnsatgastcysmetcl	562
D	2055	ACTGCATGAACACAAAGCTTTTAAAGCTTGTGATTAAGCAACACAGCAACTATTACAG	2114
OY	562	uSerphenleuclualalalaserleuemetserglnvalsertryprohlsleuvalleucl	582
D	2115	GTTCTTTCTTTGAAGCACACAGATGATGAGCAAGCTTCTCAAGACATTTGGTTTAA	2174
OY	582	uhslglylvalcysmetkilaclasp---serllmetvalglnluphevaltryleucl	601
D	2175	TTATGAGATATGTCTGCTGTGAGACAGAAATATTTCTGTTCAGGAGTTTGTAAATTTGG	2234
OY	601	yalalalaspmeteytryleuaurarglyasrarglyhnsleuvalseralasertrylsleucl	621
D	2235	ATCACTGTGATCATATCTGTGAAAAGAAATAAATTTGTATTAATATATTATGGAACCTTGA	2294
OY	621	nvalthrlysglnleuallatryalaleuasntryleuclinsaplysglyleuoprohlsgl	641
D	2295	AGTTGTCAAAACAGTTGGCAATGGCCATTCGATTTTCAAGAAAGAAACACCCATTATTCATGG	2354
OY	641	yaenvalaserlalarqyvalaleuileuulalaroglnuglycys-----asrglyasnrpr	659
D	2355	GAAATGTATGTGCCAAAATATTTCTGCTTATCAGACAGAAAGAACACAGAAATTC	2414
OY	659	opropheleuyleuenserapprogllyalserprothryalaleuenserleuclmetle	679
D	2415	TCTCTTTCATCAAACTTATAGTATCTGTGCAATTTGATTAACAGTTTGGCAAAAGACATCT	2474
OY	679	uthrasparpaleprothrvalalarproglucysleuclnglnlualaglnthrleucysle	699
D	2475	TCAGAGAGAAATACATGGGTACCACCTGAATAGCATGAAATCTTAATAATTTTAAATTT	2534
OY	699	uglnlualaspystrpelyrhecllyalathrthrtryrglnuvalpheglnatrglyproal	719
D	2535	GGCAACAGCAAAATGAGATTTTGGTACACACTTTGGGGAATCTGCACATGGAGGAGATA	2594
OY	719	ahlslllethserleuclurproalalayslysluyspheytrygluasrpglnlgnle	739
D	2595	ACCTCTTAAGTGTCTGATTTCTCAAAAGAGGTACAAATTTTATGAAGATTAAGCACTGACGT	2654

QY	739	uProAlaleuLystrPhrIngluLeuAlaGluLeuLeIethInGInuSmetAlaTyrAspRr	759
Db	2655	TCCTGTACCAAAAGTGTGGGAGATTAGCAAAACCTTATTAATTAATTTGATGGATTATGACAA	2714
QY	759	oGluTArGArProSerPheAlaGAlaIleLeuAlaGAsrLeuAsnGlyLeuLeIethSerAs	779
Db	2715	AGATTTCACAGCCTTTCTTCACAGCCATCATATCAAGATCTTAACAGATTGTATTATCTCCAGA	2774
QY	779	pTyGluLeuLeuSerAspProThrProGlyIleProSerProArg---AspAluLeuCy	798
Db	2775	TTATGAACTATTAAACAGAA---AATGACATGTTATACCAATATATAGAGATAGTGGTCCGGG	2831
QY	798	sValAlaGluAlaGlnLeuTyrAlaCysGlnAspRroAlaIlePheGlnGluAlaGlnIstLe	818
Db	2832	GTTTTCTGTGGTGC-----TTTGAAGACCGGGATCTTCACACACTTTTGAAGACAGACATT	2885
QY	818	uLeSTyTIlleSerLeuLeuGluGlyLysGlyAsnPheGlySerValaGlnLeuSuaArgTyrAs	838
Db	2886	GAATTTCTACAGCAACTTGGSCAGGGTATTTTTGGAGTGTGGAGATATGCTCCGGTATGCA	2945
QY	838	ProLeuLeuLysAspAsnThrGlyProLeuValAlaValLysGlnLeuGlnIstSerValPr	858
Db	2946	CCCTGTACAGCAAAACATGGGGAGGTGGTGGTGTAAAAAAGCTTCACACATATATACGCA	3005
QY	858	oAspRInGlnArGAspPheGlnAlaArgGlyIleGlnIleLeuLysValaLeuIstSerAspRr	878
Db	3006	ACAGCACTAAGAGACACTTTGAAAGAGAAATTTGAATCTGMAATCCCTACACATAGACAA	3065
QY	878	eIleValLysTyrArgGlyValaSerTyrGlyProGlyIleArgGlnSerLeuAlaGluValaI	898
Db	3066	CATTGTAAAGATACAGAGAGAGTGTCTACACATGCTGGTCCGGCTAATCTTAAATTAATAT	3125
QY	898	tGluTyrLeuLeuProSerGlyCysLeuAlaArgAsrLeuGlnIstAsnGlnIstArg---GlyLeu	917
Db	3126	GGAATATTATACCATATGGAAGTTTACAGACCTATTCTTCAAAAACATTAAGAACGGATAGA	3185
QY	917	sThrAspArgLeuLeuLeuPheAlaIleArgGlnIleCysLysGlyMetGluTyrLeuGlyAl	937
Db	3186	TCACATTAACACTCTGTGCAGATACACACATCTGCAGATATGCAAGGGTATGGAGTATCTTGGTAC	3245
QY	937	ArArgArgGlyValaHisArgAsrLeuLeuAlaAlaArgAsnIleLeuValaGlnSerGlyAla	957
Db	3246	AAAAAGGATATATCCACAGGGATCTGGCAACGCAAAATTAATTTGGTGGGAACCGAAACAG	3305
QY	957	sValLysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAspTyrTyrVal	977
Db	3306	AGTTAAATATTGGRGATTTTGGGTTAACCAAAAGCTTGGCCACAAAGCAAAAGATACTATTA	3365
QY	977	lValArgGluProGlyGlnSerProIlePheTrrTyrAlaProGlyIstSerLeuSerAspAs	997
Db	3366	AGTAAAGAACCTGTGTCAAAAGTCCCATATTTCTGTGTATCTCCAGAAATCACTAGCAGAGAG	3425
QY	997	nIlePheSerArgGlnSerAsrValaIlePheSerPheGlyValaValaLeuTyrGlnLeuPheTh	1017
Db	3426	CAAGTTTCTGTGGCCCTCAATATCTTGGAGCTTTGGAGTGGTCTTGATATGAACCTTTTTCAC	3485
QY	1017	rTyCysAspLysSerCysSerProSerAlaGlnPheLeuArgMetGluGlyProGlyAla	1037
Db	3486	ATACATTAGAGAAAGGTAAAGTGTCCACACACGGAATTTATGCTATGATTCAGCATATGACAA	3545
QY	1037	gGluGlyProProLeuLys---ArgLeuLeuGlnLeuLeuAlaGlnGluAlaArgAlaGlyLeu	1056
Db	3546	ACAAGGACAGATATGCTGTTCATTTGATATAGAACTTTTGAAGAATTAAGAAATTAATAC	3605
QY	1056	oProProThrCysProThrGlnValaGlnGluLeuMetGlnLeuCysTrrAlaProGly	1076
Db	3606	AAGACACGATGATGCCACGATAGATCTATGTATGATCATGACAGAAATGCTGGACATTA	3665
QY	1076	uProHisAspArgProAlaPheLeuIleThrLeuSerProGlnLeuAspProLeu	1093
Db	3666	TGTAAATCAACAGCCCTCTTAAGGATATACCTCTTGTGAGATGTGCAAAATA	3717

RESULT 6	
AAD24311	
ID	AAD24311 standard; DNA; 5117 BP.
XX	
AC	AAD24311:
XX	
DT	07-MAR-2002 (first entry)
XX	
DE	Human Jak2 (Janus kinase) DNA.
XX	
KW	Human; SOCS; suppressor of cytokine signalling; autoimmune disorder; Jak;
KW	Janus kinase; signal transducer and activator of transcription; Stat;
KW	transplant acceptance; systemic lupus erythematosus; glomerulonephritis;
KW	reumatoid arthritis; Wegener's granulomatosis; chronic active hepatitis;
KW	atopic dermatitis; multiple sclerosis; myasthenia gravis; vasculitis;
KW	Crohn's disease; haemolytic anaemia; nephrotic syndrome; dermatological;
KW	diabetes mellitus; thyroiditis; inflammatory bowel disease; nephrotropic;
KW	immunosuppressive; antiinflammatory; antirheumatic; antiarthritic;
KW	hepatotropic; ds.
XX	
OS	Homo sapiens.
XX	
FH	Location/Qualifiers
FT	478..3876
FT	CDS
FT	/*tag= a
FT	/product= "Human Jak2 protein"
XX	
PN	MO200179555-A2.
XX	
PD	25-OCT-2001.
XX	
PF	13-APR-2001; 2001WO-US12131.
XX	
PR	14-APR-2000; 2000US-0549654.
XX	
PA	(MILL-) MILENNIUM PHARM INC.
XX	
PI	Hancock WM, Ozkaynak E;
DR	WPI; 2002-034368/04.
XX	
DR	P-PSDB; AAEL15179.
XX	
PT	Monitoring transplant acceptance or autoimmune disease, useful e.g. for
PT	assessing therapy, comprises measuring levels of Stat or their
PT	inhibitors -
XX	
PS	Example; Fig 8; 21bpp; English.
XX	
CC	The present invention relates to a method for monitoring acceptance of a
CC	transplant or an autoimmune disease in a mammal. The method comprising
CC	determining the amount of at least one of Stat4 (signal transducer and
CC	activator of transcription), Stat6, SOCS1 (suppressor of cytokine
CC	signalling) or SOCS2 mRNA or protein in a sample of the transplant, take
CC	from the host or an affected tissue sample. Stats are activated by
CC	receptor-associated Janus kinases (Jaks) which include Jak1, Jak2, Tyk2,
CC	Jak3. The method is used to determine whether acceptance of a transplant
CC	has been induced or to determine if autoimmune disorders (systemic
CC	lupus erythematosus, glomerulonephritis, rheumatoid arthritis, Wegener's
CC	granulomatosis, chronic active hepatitis, atopic dermatitis, multiple
CC	sclerosis, myasthenia gravis, haemolytic anaemia, nephrotic syndrome,
CC	thyroiditis, diabetes mellitus, Crohn's disease, inflammatory bowel
CC	disease and vasculitis are being treated successfully and may be used
CC	to adjust treatment regimes. The present sequence is human Jak2 DNA.
XX	
SQ	Sequence 5117 BP; 1623 A; 927 C; 1103 G; 1464 T; 0 other;
XX	
Alignment Scores:	
Pred. No.:	2,84e-190
Score:	2621.00
Percent Similarity:	66.12%
Best Local Similarity:	48.54%
Query Match:	44.73%
DB:	24
	Gaps: 14
	5117
	Matches: 533
	Conservative: 193
	Mismatches: 341
	Indels: 31
	14

US-09-397-967-16 (1-1099) x AAD24311(1-5117)

0y	20	SeISerGuaIaGlyAlaLeuHISValILLeuAspPro-ProtArgGlyProGlyProProGly	39
0y	20	SeISerGuaIaGlyAlaLeuHISValILLeuAspPro-ProtArgGlyProGlyProProGly	39
Db	585	TTCCAGTTCTTCAGGT-----GTATCTTTACCAATTCCTTGGGAATGTAGGCAGA	635
0y	39	nArgLeuSerPheSerPheGlyAspIrrLeuAlaGluAspLeuCyValaArgAlaIaLy	59
Db	636	TTATCTGACCTTTCCATCTGGGAGATATGTGCAGAGAAGAAATCTGTATGCTGCTTAA	695
0y	59	sAlaCySgIyLLeuProValIyTHISerLeuPheAlaLeuAlaTHnGluAspPhe	79
Db	696	AGCTTGAGTTCACACCTGTGTATATATATATGTTTGTATTAATAGCTGAACAGAAAG	755
0y	79	rCyStrPheProPheSerHISLeuPheCySgIeGluAspValaSPnGlnValLeuVa	99
Db	756	GATCGGTATCCACCACCAATGCTTCCATATAGATGAGTACACAGGCAATATGTCT	815
0y	99	LYrArGLeuArGpHeTyRrPheProAsPTrPhe-----GlyLeuGluTHrCySHsAr	117
Db	816	CTACAGATATAGATTTTACTTCTCTGTTGATGTGCAGGACACCAAGACCTATAG	875
0y	117	gPheGlyLeuArGlyAspLeuThrSerAlaILLeuAspLeuHISValLeuGluHIS	137
Db	876	GCATGGAATATCTGAGGTGCTGAGGTCTCTCTTGTGATGACTTGTGCATGTCTTACT	935
0y	137	uPheAlaGlnHISrGSerAspLeuValSerGlyArgLeuProValGlyLeuSerMeLy	157
Db	936	CTTCTCTAGTGGCGGCATGATTTTGTGCAGGATGATTAAGTAACGTGATCTATGA	995
0y	157	sGluGlnGlyLubPheLeuSerLeuAlaValLeuAspLeuAlaGlnMeAlaArgGlu	177
Db	996	AACACAGAAAGAAATGCTCTGGATGACGACGTGTAGATATGATGATGACAAAGAAA	1055
0y	177	nAlaGlnArgProGlyLLeuLeuLeuLySThrValSerTyLISAlaCyLeuProProSe	197
Db	1056	CGATCAAAACCCACTGGCCATCTTAACCTGTATACAGCTACAGACACTTGTTCACAAATG	1115
0y	197	rLeuArGAspValILInGlnGluAsnPheValThrArgArgArgILeArgTrHVa	217
Db	1116	TATGTGACCAAGAAATCCAGACACTCATATATTTGCACAGAAAGCAATGAGGTACAGAT	1175
0y	217	lValLeuAlaLeuLeuProCySgIArgLeuProGlyArgProTyAlaLeuMeAlaLy	237
Db	1176	TGCGAATTTATATAGCAATTCAGCAATTCAGCAATTCAGCAATTCAGCAATTCAGCAAT	1235
0y	237	sTyxILLeuAspLeuGluArgLeuHISProAlaAlaThrThrGluThrPheArgValG	257
Db	1236	GTAATCTTAATAATCTGGAACCTGCGAGTCTGCGCTTACACAGAGAATTTGAAGTAA	1295
0y	257	yLeuProGlyAla-----GlnGluLubProGlyLeuLeuArgValAlaGlyAs	273
Db	1296	AGAACCTGGAATGCTGCTCAGGTGAGAGAGATTTTTCAGCAATCATATATATACGTGAA	1355
0y	273	pAsnGlyLLeuProTrpSerSer-----AsnAspGluLe	284
Db	1356	CGGTGGAATTCAGTGTGTCACAGAGGGAACATTAAGAAAGTGAACACTGACAGAGAGA	1415
0y	284	uPheGlnThrPheCySAspPheProGluILLeValaSPValSerILAsnGlnAlaProAr	304
Db	1416	TTTACAGTTATATTCGATTTTCTCAATATATATGATGATCAGATTAAGACCAACCA	1475
0y	304	gValaGlyProAlaGluHISArgLeuValThrValThrArgMeLAspGlyHISLele	324
Db	1476	AGAGGCT--TCAAATGCAAGCCGAGTGTAACTATCCATTAAGAGATGCTGTAATAATCT	153
0y	324	uGluAlaGluPheProGlyLeuProGluAlaLeuSerPheValAlaLeuValaSPGlyLy	344
Db	1533	GGAAATTTGAACCTATCAATTAAGGAGAGCTTGTCTGTGTCATTAATATATGATGA	159
0y	344	rPheArgLeuILLeuCyAspSerArgHISTyRPhCySgIySglnValaLabProbraGle	364

Db 1593 TTATACATTAACTGCAGATGCACATCATTAACCTCTGTAAAGAGTAGACCTCCAGCCGT 1652
 Qy uleuGluGluGluAlaAspValCysHisGlyProIleThrLeuAspPheAlaIleHisLy 384
 Db 1653 GCTTGAATAATATACAAAGCACTGTGATGGCCCAATTTGCATGGATTTGGCCATTGTGTA 1712
 Qy sleuLysAlaIaGlySerLeuProGlyThrTyrIleLeuArgArgSerProGlnAsp 404
 Db 1713 ACTGAAGAAAGCAGATATACAGACGTGACGTATGTACTGATGCATGACGCTTAAGACATT 1772
 Qy rasSerPheLeuLeuThrAlaCysValGlnThrProLeuGlyProAspTyrIleGly 424
 Db 1773 TAATTAATATTTTGTGATCTTGTGCTGTGCGAGAAATGTCTGATTAATATTAACACATG 1832
 Qy sleuIleArgGlnAspSerSerGlyAlaPheSerLeuValGlyLeuSerGlnProHisAr 444
 Db 1833 TTTCGATTACAAAATAAGACATGAATGAATGACAACTCAGTGGACAAAGAACTTCAG 1892
 Qy gSerLeuArgGluLeuLeuAlaIaCysThrPasnSerGlyLeuArgValAspIleAla 464
 Db 1893 CAGTCTTAAAGATCTTTGAAATGTTCACAGATGGAACCTGTGCTCAGACAAATTAAT 1952
 Qy aleuTyrLeuThrSerCysCysAlaProAlaArgProLysGluLysSerAsnLeuIleVal 484
 Db 1953 TTTCAGATTACTAAATGCTGTCCCAAGCCAAAGATCAAACTCTCTAGCTCTT 2012
 Qy 484 IArg---ArgGlyCysAsnProAlaProAlaProGlyCysSerProSerCysAlaLe 503
 Db 2013 CAGAACGATGCTGTTTGCATGTACCAACCTCACAACATTCACAGGCTCTCATAT 2072
 Qy 503 uThrGlnLeuSerPheHisThrIleProThrAspSerLeuGluTyrPheHisGlnAsnLeu 523
 Db 2073 GAACCAAAAGTGTTTACAAAATCAGAAATCAGAAATGATGATTAATGAAGACCTTGG 2132
 Qy 523 yHisGlySerPheThrLysIlePheArgLysSerArgArgValAlaAsp---GlyLeu 542
 Db 2133 CCAGGCACTTTTACAAAGATTTTAAAGCCGTACAGAGAGATGAGAGACAGCTGTA 2192
 Qy 542 uThrHisAspSerGluValLeuLeuLysValIleCAspSerArgHisArgAsnCysMetG 562
 Db 2193 ACTGCATGAACAGAGATCTTTTAAAGTTCGTGATTAAGCAGACAGAACTATTCAGA 2252
 Qy 562 uSerPheLeuGluAlaIaIleSerLeuMetSerGlnValSerTyrProHisLeuValLeu 582
 Db 2253 GCTCTTCTTGAAGCAGAGATGATGATGAGCAAGCTTCTCACACACATTGGTTTAA 2312
 Qy 582 uHisGlyValCysMetAlaGlyAsp---SerIleMetValGlnGluPheValTyrLeuG 601
 Db 2313 TTATGCAATATGCTCTGTGAGACAGACGAATATCTGCTTCAGAGATTGTGTAATTTGG 2372
 Qy 601 yAlaIleAspMetTyrLeuArgLysArgGlyHisLeuValSerAlaSerTyrPheLysLeuG 621
 Db 2373 ATCAGTACGATACATCTGAAAGAAATGAATTAATGTATATGGAACCTGA 2432
 Qy 621 nValThrLysGlnLeuAlaIleTyrAlaLeuAsnTyrIleGluLysAspGlyLeuProHisG 641
 Db 2433 AGTTGCTAAACAGTTGGCCATGGCCATGCAATTTCTAGAGAAACACCCCTTAATGATG 2492
 Qy 641 yAsnValSerAlaArgLysValLeuLeuAlaIArgGluGlyLy-----AspGlyAsnPr 659
 Db 2493 GAATGATATGCGCAAAATAATCTGCTTATCAGAGAAAGACAGACAGAAACAGAAATCC 2552
 Qy 659 oProPheIleLysLeuSerAspProGlyValSerProThrValLeuSerLeuGluMetLe 679
 Db 2553 TCCCTTCATCAAACTAGATCTGTGCAATAGTATTAACAGTTTCCCAAGACATTCCT 2612
 Qy 679 uThrAspArgIleProThrValAlaProGluCysLeuGlnGluAlaGlnThrLeuCysLe 699
 Db 2613 TCAGAGAGAAATACATGGGTACACCTGAATGATGCAATGCAATTCCTAAAAATTTAAATTT 2672
 Qy 699 uGluLysAspLysTyrPglYpHegIyAlaIleThrThrTrpGluValPheGlnArgGlyProAl 719
 Db 2673 GGCAACAGACAAATGAGAGTTTGGTACCACTTTGTGGGAATCTGACAGTGGAGAGATTA 2732

Qy 719 aHisIleThrSerLeuGluProAlaLysLysLeuLysPheTyrGluAspGlnGlyIle 739
 Db 2733 ACCTTAAGTGTCTGTGATTTCTCAAGAAAGATACAAATTTTATGAAGATGAGATGAT 2792
 Qy 739 uProAlaLeuLysTyrTrpThrGluLeuAlaGlyLeuIleThrGlnCysMetAlaTyrAsp 759
 Db 2793 TCCGTGACCAAAAGTGGCGAGAAATTACCAACCTTATTAATTAATGTATGATTAATGACC 2852
 Qy 759 oGlyArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIleThrSerAs 779
 Db 2853 AGATTTACAGCCCTTCTTACAGCCCATATACAGATCTTAAACAGTTTGTATCTCACA 2912
 Qy 779 pTyrGluLeuLeuSerAspProThrProGlyIleThrProSerProArg---AspGlyLeu 798
 Db 2913 TTATGCACTATTACAGAA---AATGACATGTTTACCAAAATGAGAGATGAGTGGCCCTGG 2969
 Qy 798 sValAlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGluGluArgHisLe 818
 Db 2970 GTTTTCTGGTGC-----TTTGAAGACCGGATCCTACAGATTTGAAGAGACATTT 3023
 Qy 818 uLysTyrIleSerLeuLeuGluLysGluLysAsnProGlySerValGlnLeuCysArgTyrAs 838
 Db 3024 GAATTTCTACACAACTTGGCAAGGTAATTTTGGAGTGTGAGATGTGCGGCTATGA 3083
 Qy 838 pProLeuGlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGlnHisSerValPr 858
 Db 3084 CCGTCTACAGACAAACACTGGGAGGTGCTGCTGTAACAAACCTTCACACTATGCTGA 3143
 Qy 858 oAspGlnGlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuHisSerAspPh 878
 Db 3144 AGAGCACTTAAGCACTTTGAAGAGGAATTTGAATCTCGAAATCCCTCAGACATGACA 3203
 Qy 878 eIleValLysTyrArgGlyValSerTyrGluProGlyAlaArgGlnSerLeuArgValLe 898
 Db 3204 CATGTAAAGTCAAGAGAGTGTGTACAGTGTGCTGCTGCGGCTGAATCTGAATTAATTA 3263
 Qy 898 tGluTyrLeuProSerGlyCysLeuArgAspLeuGlnArgHisArg---GlyLeuHis 917
 Db 3264 GGAATTTTACCATATGAGAAATTTACAGACATATCTTCAAAAACATTAAGAACGATGA 3323
 Qy 917 sThrAspArgLeuLeuPheAlaIleTrpIleCysLysGlyMetGluTyrLeuGlyAl 937
 Db 3324 TCACATTAACACTTCCATCCATCATCATCATCAAGATTCACAAAGGATATGAGTATCTTG 3383
 Qy 937 aArgArgCysValHisArgAspLeuAlaIleArgAsnIleLeuValGlnSerGluLan 957
 Db 3384 AAAAGGCTATTCACAGAGATCTGCAACAGAAATATATGTTGGTGGACAGACAGACAG 3443
 Qy 957 sValLysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAspTyrTyrVa 977
 Db 3444 AGTTAAATTTGAGATTTGGGTTAACCAAGCTTGCCACAAAGCAAAAGATATGATTA 3503
 Qy 977 lValArgGluProGlyGlnSerProIlePheTrpTyrAlaProGluSerLeuSerAspAs 997
 Db 3504 AGTAAAGAACCTGTGGAAAGTCCCATATTTCTGTGATCTCCAGAAATTCACAGACAG 3563
 Qy 997 nIlePheSerArgGlnSerAspValTyrSerPheGlyValAlaIleTyrGluLeuPheTh 1017
 Db 3564 CAAGTTTCTGTGGCTCAGATGTTTGGAGCTTTGGAGTGTCTGTATGAACCTTTTCA 3623
 Qy 1017 rTyrCysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetGlyProGlu 1037
 Db 3624 ATACATTTGAGAGATGAAGATCCACACCGGAATTTATGCTATGATTTGGCAATGACAA 3683
 Qy 1037 gGluGlyProProLeuCys---ArgLeuLeuGluLeuLeuAlaGluGlyArgArgLeuPr 1056
 Db 3684 ACAAGACAGATGATGCTGCTCCATTTGATTAACATTTTGAAGAAATATGAGAAATTAAC 3743
 Qy 1056 oProProProThrCysProThrGluValGlnIleLeuMetGlnLeuCysTyrAlaProGl 1076
 Db 3744 AAGACCAAGATGATGCCAGATGATGATATATGATCATGACAGAAATGCTGGAACAAATTA 3803

OY 1076 uProHisAspArgProAlaPheAlaThrLeuSerProGlnLeuAspProLeu 1093
DB 3804 TGTAAATCAACGCCCTTACGATCTAGCTTCTGAGTCAATA 3855
RESULT 7
AA085412
ID AA085412 standard; cDNA; 3629 BP.
XX
AC AA085412;
XX
DT 05-OCT-1995 (first entry)
XX
DE Murine JAK2 kinase coding sequence and flanking regions.
XX
KW JAK family; protein tyrosine kinase; cytokine receptor; mouse;
KW phosphorylation; signal transduction; activation; ss.
OS
XX Mus musculus.
FH
FT Key Location/Qualifiers
FT CDS 94..3483
FT
FT /*tag= a
FT /*product= Jak2_kinase
FT 554..556
FT conflict /*tag= b
FT /*note= "Published partial sequence of Jak2 cDNA
FT (Harpur et al., Oncogene 7:1347-1353(1992))
FT has CCC at this position"
FT 1089..3579
FT conflict /*tag= c
FT /*note= "Published partial sequence of Jak2 cDNA
FT (Harpur et al., Oncogene 7:1347-1353(1992))
FT differs from AA085412 in having the
FT nucleotides shown in brackets at the following
FT positions: 1089(T), 1103(C), 1114(G), 1119(G),
FT 1122(C), 1128(C), 1131(G), 1134(G), 1137(C),
FT 1140(G), 1143(G), 1146(C), 1188(T), 1194(G),
FT 1230(A), 1245(T), 1260(T), 1266(C), 1272(T),
FT 1275(A), 1293(T), 1305(T), 1323(C), 1341(A),
FT 1344(A), 1359(G), 1365(A), 1368(T), 1374(T),
FT 1401(C), 1413(C), 1431(T), 1458(A), 1476(G),
FT 1488(T), 1511..1512(GT), 1578(C), 1590(T),
FT 1593(T), 1602(G), 1623(T), 1642(G), 1657(C),
FT 1728(G), 1743(C), 1755(C), 1770(A), 1809(G),
FT 1816(G), 1821(C), 1857(A), 1878(T), 1935(A),
FT 1938(A), 1963(T), 1974(G), 2025(T), 2055(G),
FT 2079(C), 2082(C), 2085(C), 2253(A), 2259(G),
FT 2283(A), 2285(C), 2433(G), 3453(C), 3579(C)."
FT conflict 2226
FT /*tag= d
FT /*note= "location of a 7 amino acid insert detected
FT in previous studies, but not in the present
FT study"
FT 3595..3619
FT conflict /*tag= e
FT /*note= "three extra nucleotides (all A's) were
FT present in previous studies at positions
FT 3595, 3598 and 3619 in the 3'-UTR"
XX
PN M09503701-A.
XX
XX 09-FEB-1995.
PD
XX
XX 29-JUL-1994; 94MO-US08676.
PF
XX
XX 29-JUL-1993; 93US-0097997.
PR
XX
XX (STUD-) ST JUDE CHILDREN'S RES HOSPITAL.
PA
XX
XX Inle JN, Quelle FW, Silvenoinen O, Witthuhn BA;
PI
XX WPL: 1995-081950/11.
DR
XX P-PSDB: AAR70830.

XX
PT Inhibiting a cellular response to a cytokine by inhibiting Jak
PT kinase - to treat diseases caused by excessive response to
PT cytokine, e.g. erythrocytosis and other cellular proliferative
PT diseases
XX
XX
PS Claim 29; Fig 1; 167pp; English.
XX
CC Inhibiting the activity of a jak kinase (pref. Jak1, Jak3 or
CC Tyk2) in a eukaryotic cell is claimed as a method of inhibiting the
CC biological response of that cell to a cytokine (not IL-3 or
CC erythropoietin). The present sequence (murine JAK2 kinase) encodes a
CC 1129 amino acid protein which includes an epitopic sequence at
CC amino acid positions 758-776. Antibodies which selectively bind the
CC epitope are able to bind Jak2 without interfering with the activity
CC of the kinase. Such antibodies are claimed and are useful for
CC detecting and extracting Jak2. The 71 nucleotide differences noted
CC between the present sequence and the published partial sequence
CC result in 9 amino acid changes.
XX
SQ Sequence 3629 BP; 1142 A; 733 C; 842 G; 912 T; 0 other;
XX
Alignment Scores:
Pred. No.: 2,33e-190 Length: 3629
Score: 2619.50 Matches: 537
Percent Similarity: 65.50% Conservative: 190
Best Local Similarity: 48.38% Mismatches: 350
Query Match: 44.70% Indels: 33
DB: 16 Gaps: 15
US-09-397-967-16 (1-1099) x AA085412 (1-3629)
OY 4 ProSerGluGluThrProLeuIleProGlnArgSerCysSerLeuSerSerGluAla 23
DB 145 CCTGTACATCAGANNGATGATTCCTGGAGGCTGTAATTCGTGAGACAGATAGACCA 204
OY 24 GlyAlaLeuHisValLeuLeuProProArgGlyProGlyProProGlnArgLeuSerPhe 43
DB 205 GTCCCTCAAGTGTATCTGTACCATTCCTCTGGCAAGCTGAAGAGATATCAAGTTT 264
OY 44 SerPheGlyAspPyrLeuAlaGluAspLeuCysValArgAlaAlaValSAlaCysGlyIle 63
DB 265 CCAAGTGGAGAGATGTGTTCACAGAGAAATTTGTGTGCTGTTTAAAGCTTGATATT 324
OY 64 LeuProValThrHisSerLeuPheAlaLeuAlaThrGlnAspPheSerCysTrpPhePro 83
DB 325 ACGGCTGGTATCAAAATATGTTGCGTTAATGAGTAAACCAAGAGATCTGACCCA 384
OY 84 ProSerHisIlePheCysIleGluAspValAspThrGlnValLeuValTyrArgLeuArg 103
DB 385 CCCAATCATGTGTCACATATGACGAGTCAACAGGATGACATATCTTACAGATTAAGG 444
OY 104 PheTyrPheProAspTrpPhe-----GlyLeuGluThrCysHisAspPheGlyLeuArg 121
DB 445 TTCTACTTCCTCATGTGTTAGTGTAGTGCAGCAGCAACCACTTACATAGAGGTGCC 504
OY 122 LysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGlnHisLeuPheAlaGlnHis 141
DB 505 CGTGCGGCTGAAGCTCCCTGCTGATGACTTGTGTCATGCTTACCTTTTGTCTCAGTGG 564
OY 142 ArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLysGluGlnGly 161
DB 565 CGGCATGATTTTGTTCACGAGTGAATAAAGTACCTGTGACTCATGAACTCAGAGAGAG 624
OY 162 PheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGlnAlaGlnArgPro 181
DB 625 TGTCTGGGATGCGCGTGTAGACATGATGAGANTAGTAAAGAGAAAGACAGACTCCA 684
OY 182 GlyIleLeuLeuLysThrValSerTyrLysAlaCysLeuProProSerLeuArgAspVal 201
DB 685 CTGGCTGTCTAATCTGTGACGACACAAACATTTTACCAAGAGTGTGCGAGCGAG 744
OY 202 IleGlnGlyGlnAsnPheValThrArgArgArgIleArgArgIleValValLeuAlaLeu 221

```

Db 745 ATCCAAAGACTATACATTTTAAACCCGGAAGCGAATCAGATTCAGATTCAGATTCATT 804
Oy 222 LeuProGlyValLeuLeuProGlyArgProTyrAlaLeuMetAlaLysTyrIleLeuAsp 241
Db 805 CAGCAATTACGTCATGTAAGAACCCACTGCCAGAACCTTAACCTTAAGATTCATTAAAC 864
Oy 242 LeuGluArgLeuHisProAlaIleThrThrGluThrPheArgVal 256
Db 865 CTGGAAACCCCTGACGTCGCTCTACACAGAACGTTTGAAGTAAGAATCTGCAGAA 924
Oy 257 GlyLeuProGlyAlaGlnGluGluProGlyLeuLeuArgValAlaGlyAspAsnGlyIle 276
Db 925 GGT---CTTTCAAGGTGAGAGATTTTTCACACCATTTAATTAAGTGAACCGTGAAT 981
Oy 277 ProTyrSerSer 287
Db 982 CAGTGTCTCAAGGGAACATAAGAAAGTGAACACTGACACAGAACGAGACGTA 1041
Oy 288 PheCysAspPheProGluIleValAspValSerIleAsnGlnAlaProArgValGlyPro 307
Db 1042 TATTGTGATTTCCCTGATATTTATGATGTACGATTAATTAAGCAAGAACACG---GAATGC 1098
Oy 308 AlaGlyGlnHisArgLeuValThrValIleThrArgMetAspGlyHisIleLeuGluAlaGlu 327
Db 1099 TCAAATGAAGTAGAATGTAATGTAAGTCTGCTAAACAAGATGTAAGTTTGGAGATAGA 1158
Oy 328 PheProGlyLeuProGlyAlaLeuSerPheValAlaLeuValAspGlyTyrPheArgLeu 347
Db 1159 CTTAGCTGATTAAGAAGACCTTGTGATTCGTCATTAATTAAGCGGATTAACAGACTA 1218
Oy 348 IleCysAspSerArgHisTyrPheCysGlyGluValAlaProArgLeuLeuGluGlu 367
Db 1219 ACTGGGAGTCCACCATCTGCTCAAGAGAGTGGCCCCCGAGCTGTGCTCGAGAAC 1278
Oy 368 GluAlaAspValCysHisGlyProIleThrLeuAspPheAlaIleHisIleGlyLeuAla 387
Db 1279 ATACACAGCAACTGCCACGCCCAATATGATGATTTGGCATTAAGCAACTTAAGAAG 1338
Oy 388 AlaGlySerLeuProGlyIleThrTyrIleLeuArgArgSerProGlyAspTyrAspSerPhe 407
Db 1339 GCGGCTAACACAGACTGACTATATGCTACAGTACGAGCCCTAAGAGACTTCAACAATAC 1398
Oy 408 LeuLeuThrAlaCysValAlaIleThrProLeuGlyProAspTyrGlyGlyCysLeuIleArg 427
Db 1399 TTTTGTACCTTTGCTGTTGAGCGAGAAATGTCAATTAATTAACACCTTTGATTCAG 1458
Oy 428 GlnAspProSerGlyAlaPheSerLeuValGlyLeuSerGlnProHisArgSerLeuArg 447
Db 1459 AAGATGAGAAATGAGCAATACACTACGCGGACTAAGAGAACTTCAGTAACCTTAAG 1518
Oy 448 GluLeuLeuAlaAlaCysTyrPheAsnSerGlyLeuArgValAspGlyAlaAlaLeuTyrLeu 467
Db 1519 GACCTTTTGAATTCCTACACAGATGGAACCTGCGCTCAGACAGATCATCTTCCAGTTT 1578
Oy 468 ThrSerCysCysAlaProArgProGlyGluCysSerAsnLeuIleValAlaIleArgGly 487
Db 1579 ACCAAATGCTGCCCCCAAGCAAAAGATAATCAAACTTCTGCTCAGAG--- 1632
Oy 488 CysAsnProAlaProAlaProGlyCysSerProSerCysAla 504
Db 1633 ACAATGTGATTTCTGATGTTCCAGATCTCACCAACATTAACAGAGCAATTAATGAT 1692
Oy 505 GlnLeuSerPheHisThrIleProThrAspSerLeuGluTyrPheHisGluAsnLeuGlyHis 524
Db 1693 CAAATGGTGTTCACAAATACGAAATGAGATTAATTAATTAAGAAAGCTTGGCCAA 1752
Oy 525 GlySerPheThrIlePheArgGlySerArgGlyValAlaAsp---GlyLeuThr 543
Db 1753 GGTACTTTTACAAATTTTAAAGGTGTAAGAAAGAGATTTGGAGATTTAGTGAACCTG 1812
Oy 544 HisAspSerGluValLeuLeuLysValIleAspSerSerArgHisArgAsnCysMetGluSer 563

```

```

Db 1813 CACAAAGCGAAGTCTTTTGAAGTCTAGATTAAGCATAGCACTATTCAGAGTCT 1872
Oy 564 PheLeuGluAlaAlaSerLeuMetSerGlnValSerTyrProHisLeuValLeuLeuHis 583
Db 1873 TTTCTTGAAGCAGCAAGCATGATGATGATGATGATGATGATGATGATGATGATGAT 1932
Oy 584 GlyValCysMetAlaGly---AspSerIleMetValGlnGluPheValTyrLeuGlyAla 602
Db 1933 GGTGCTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1992
Oy 603 IleAspMetTyrLeuArgLeuGlyArgGlyHisLeuValSerAlaSerTyrPheLeuGlnVal 622
Db 1993 CTGGATACATACCTGAGAAAGCAAAATTCATTAATATATTTATGGAACCTTGAGAG 2052
Oy 623 ThrLysGlnLeuAlaTyrAlaLeuAsnTyrLeuGluAspGlyGlyLeuProHisGlyAsn 642
Db 2053 GCTAAGCACTGTGGCATGGCCATGCTTTTGAAGAAATAATCCCTTAATCATGAGGAT 2112
Oy 643 ValSerAlaArgValLeuLeuValAlaArgGluGlyGly-----AspGlyAsnProPro 660
Db 2113 GTGTGTGCTTAAATATATCTGCTTATACAGATGAAGAAAGACAGAGAACGGGAACTT 2172
Oy 661 PheIleLysLeuSerAspProGlyValSerProThrValLeuSerLeuGluMetLeuThr 680
Db 2173 TTTATCAAACTTAGATGATCTGTGATGATGATGATGATGATGATGATGATGATGAT 2232
Oy 681 AspArgIleProTyrValAlaProGlyCysLeuGlnGluAlaGlnThrLeuCysLeuGlu 700
Db 2233 GAGAAATATCCATGGGTACCTCTGTGAATGCAATGCAATCAAAATCTCAATGTGCA 2292
Oy 701 AlaAspLysTyrGlyPheGlyAlaThrThrThrGluValPheGluArgGlyProAlaHis 720
Db 2293 ACAGACAAGTGAAGCTTGGGACCACTGTGTGGAGATCTGACAGTGAAGATTAAGCC 2352
Oy 721 IleThrSerLeuGluProAlaLysLysLeuLysPheTyrGluAspGlnGlyLeuPro 740
Db 2353 CTGAGTGTCTGTGATTCGAAAGAGTGCAGTTCATGAAGATTAACATCACTGCTTCT 2412
Oy 741 AlaLeuLysTyrThrGluLeuAlaGlyLeuIleThrGlnCysMetAlaTyrAspProGly 760
Db 2413 GCACCCAGTGTGACAGATTAACCAACTTATTAATTTGCATGAGCTATGAGCCACAT 2472
Oy 761 ArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIleThrSerAspTyr 780
Db 2473 TTTCAAGGCTGCTTTCAGAGCTGTATCCGATCTTAACAGCGCTGTTACTGCAAGAT 2532
Oy 781 GluLeuLeuSerAspProThrProGlyIleProSerProArg---AspGlyLeuCysVal 799
Db 2533 GAACTACTTAACAGAA---AATGACATGCTACCAACATGAGAAATAGTGCCTAGGGTTT 2589
Oy 800 AlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGluGluArgHisLeuLys 819
Db 2590 TCTGTGTCT-----TTTGAAGACAGGAGCCCTACACAGTTTGAAGAGACACTTGAAG 2643
Oy 820 TyrIleSerLeuLeuGlyLysGlyAsnPheGlySerValGlyLeuCysArgTyrAspPro 839
Db 2644 TTTCTACAGCACTTGGCAAGAGTAACTTCGGAGTGTGGAGATGTGCCCTATGAGCCG 2703
Oy 840 LeuGlyAspAsnThrGlyProLeuValAlaValLysGlnIleGlnHisSerValProAsp 859
Db 2704 CTGACAGCAACACTGCGGAGAGTGTGCTGTGAAGAAATCTCCACAGACAGTGAAGAG 2763
Oy 860 GlnGlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuHisSerAspPheIle 879
Db 2764 CACCTCCGAGACTTTGAAGAGAGATGATGATGATGATGATGATGATGATGATGATGAT 2823
Oy 880 ValLysTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuValGluValMetCys 899
Db 2824 GTCAAGTACAGAGGAGTGTCTACAGTCCGGGTGGGCCAACSTTAAGATTAATATGAAA 2883
Oy 900 TyrLeuProSerGlyCysLeuArgAspLeuLeuArgGlnArgHisIleArg---GlyLeuHisThr 918
Db 2884 TATTTACCATTTGGAAGTTTACGAGACTATCTCCAAACATTAAGAAAGGATGATGATC 2943

```

Qy	919	AsparagineuLeuPhealatrprglnileCysLysgLMetGluTyrLeuGlyAlaArg	938
Db	2944	AAAAAACTTTTAAATACACATCTCGATATGACAGGGCAGTGAATCTTGGTACAAA	3003
Qy	939	ArgCysValHisArgaspLeuAlaAlaArgasnileLeuValGluSerGluAlaHisVal	958
Db	3004	AGGTATATCCACAGGGACCTGGCACAACAGACATATTGGGAAAATAAGAACAGGGGT	3063
Qy	959	LysIleAlasphPheGlyLeuAlaLysIleLeuProLeuGlyLysAspTyrTyrVal	978
Db	3064	AAATATAGCAACTTCGGATTAACCAAAAGTCCTGGCCAGACAAAGAAATGAAACAGGGTT	3123
Qy	979	ArgGluProGlyGlnSerProIlePheTyrTyrAlaProGluSerLeuSerAspAsnIle	998
Db	3124	AAGGAGCGACGGGAAAGCCCATATCTGTGTACGACCGCTGAATCCTTGACGAGAGCA	3183
Qy	999	PheSerArgGlnSerAspValTrpSerPheGlyValValLeuTyrGluLeuPheThrTyr	1018
Db	3184	TTTTCTGTGGCCCTCAGTGTGTGTGGACCTTGGAGTCCTCTTAAGAACTTTTGACATAC	3243
Qy	1019	CysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetGlyProGluArgGlu	1038
Db	3244	ATCCAGAGAGAGTAAAGTCCACCCCTGGGAATTTATGCCAGATGATTGGCAATGATAACAA	3303
Qy	1039	GlyProProLeuCys--ArgLeuLeuGluLeuLeuAlaGluGlyArgArgLeuPro	1057
Db	3304	GGCCAAATAGATGTGTCCATTTGGATATAGACTACTGAAAGCAACGGAAGATTGGCAAG	3363
Qy	1058	ProProThrCysProThrGlnValGlnGluLeuMetGlnLeuCysTrpAlaProGluPro	1077
Db	3364	CCAAAGAGATGCCACAGATGAGATTATGTGATCATGTACAGAGTGTGGAAACAACAATGTG	3423
Qy	1078	HisAspArgProAlaPheAlaThrLeuSer	1087
Db	3424	AGCCAGCGTCCCTCTCTTCAGAGCACTTTGCG	3453
RESULT 8			
AC	66244	standard; cDNA; 3629 BP.	
AC	66244:		
DT	19-FEB-2001	(first entry)	
DE	Jak2 polynucleotide sequence.		
KW	Jak3, kinase; cytokine; cellular response; inhibition; jak2; ss;		
OS	cell proliferation; erythrocytosis.		
XX	Unidentified.		
XX	US6136595-A.		
XX	24-OCT-2000.		
XX	18-JUN-1996;	96US-0665574.	
XX	29-JUL-1994;	94US-0282012.	
XX	29-JUL-1993;	93US-0097997.	
XX	09-SEP-1993;	93US-0118968.	
XX	(SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.		
XX	Silvenoinen O, Wiltuhn BA, Ihle J;		
XX	WPI; 2000-686080/67.		
XX	P-PSDB; AAB35719.		
XX	New DNA encoding Jak3 kinase is useful as cytokine regulator for		
XX	treating cell proliferation -		
XX	Example 1; Fig 1; 100pp; English.		

XX This invention relates to DNA encoding a murine Jak3 protein. The amino
CC acid sequence of the Jak3 protein is given in AAB55715. The Jak family
CC of kinases are involved in the cellular response to the binding of
CC cytokines to their respective receptors. Jak3 kinase mediated activation
CC of some cytokines through their phosphorylation in response to
CC cytokine-receptor binding. Inhibiting the activity of Jak3 kinase (at the
CC nucleic acid level with antisense sequences or ribozymes, or at the
CC protein level with antibodies, kinase inhibitors etc.) is used to treat,
CC or diagnose, diseases caused by excessive secretion of certain cytokines,
CC e.g. excessive cell proliferation such as erythrocytosis. Alternatively,
CC the Jak3 protein and polynucleotide can be used to treat conditions
CC associated with defective Jak3 activity. The DNA sequence can be used to
CC produce recombinant Jak3 and this used to raise antibodies useful as
CC specific inhibitors or to detect or isolate Jak3 without interfering
CC with its enzymatic activity. The present sequence represents cDNA
CC encoding a Jak3 protein used in the isolation and characterisation of the
CC Jak3 protein of the invention.

XX
XX Sequence 3629 BP; 1142 A; 732 C; 843 G; 912 T; 0 other;

Alignment Scores:	2,33e-190	Length:	3629
Pred. No.:	2619.50	Matches:	537
Score:	65.508	Conservative:	190
Percent Similarity:	48.38%	Mismatches:	350
Best Local Similarity:	44.70%	Indels:	33
Query Match:	21	Gaps:	15

US-09-397-967-16 (1-1099) x AAC66244 (1-3629)
QY 4 ProsergiugluthrProleuIleProGlnargSerfScySerleuSerSerSergIuaIa 23
:::: :::: :: ::::
DB 145 CCTGACACATGCAATGGGATATTCCTCGAAGTGCCTAATTCGTGCAAGCAGATAGACCA 204
QY 24 GAlaIleuHISValLeuLeuProProArgGlyProGlyProProGlnaIgleuSerPne 43
DB 205 GTCCCTTCAGATGTCATCTGCACCATTCCTCTGGGCAAGCTGAGAGAGATCTGGAAGTTT 264
QY 44 SerPhegluVAspIlyrLeuAlaIgluAspIleucSValaIrgAlaIalrAlaIacyGlyIle 63
:::: :::: :::: :::: :::: ::::
DB 265 CCAAGTGGAGAGATGTGTGGCAGAAAGAAATTTGTGTGGCTGCTCTTAAGCTGTGGATAT 324
QY 64 LeuProValIlyrHISSerSerLeuPheAlaLeuAlaThrGluAspPheSerfSTrPhePro 83
:::: :::: :::: :::: :::: ::::
DB 325 ACGCTGTGTATCATAAATATTTGGCTTATGAGTAGAAGCAACCGAAGATCTGGTACCCA 384
QY 84 ProSerHISIlePheCysIleGluAspValaIspHrGlnaIleuValIlyrArgLeuArg 103
:::: :::: :::: :::: :::: ::::
DB 385 CCCATCATCATGCTTCCACATAGACGAGCAACCGAGCATGCATACATCTTACAGGATAGG 444
QY 104 PheIlyrPheProAspTrPhe-----GlyLeuGluThrCysHISArgPhegluLeuArg 121
:::: :::: :::: :::: :::: ::::
DB 445 TTCTACTTCCCTCATGTGTACTGACTAGTGGCAGCAGCAAGAACCTTACAGATACGGAGTCTCC 504
QY 122 LysAspLeuThrSerAlaIleLeuAspIleuHISValLeuGluHISLeuPheAlaGlnHIS 141
:::: :::: :::: :::: :::: ::::
DB 505 CGTGGGGCTGAAGCCCTCTCTCTGATCATCTTGTCAATGTCATTACCTTTTGGCTCACTGG 564
QY 142 ArgSerAspLeuValSerGlyArgLeuProValGlyLeuSermeTlYsGluGlnGly 161
:::: :::: :::: ::::
DB 565 CGGCAATGATTTTGTTCACGCGATGGATTAATACCTGTGACATCATGAACACACAGAGAG 624
QY 162 PheLeuSerIleuAlaValLeuAspLeuAlaGlnMeAlaIargIugIuAlaGlnArgPro 181
:::: :::: :::: :::: :::: ::::
DB 625 TGTCTTGGAGTGGCGGTGTAGACATGATGAGTAAGTACTAGAGGAAGAACCCAGATCCA 684
QY 182 GlyIguLeuLeuLysThrValSerIlyrLysAlaCysLeuProProSerLeuArgAspVal 201
:::: :::: :::: :::: :::: ::::
DB 685 CTGGCTGTATACCTGTGTAGCTACACACATTTTACCAAAAGTGGCTTCAGGCAAG 744
QY 202 IlegIuGlyGlnAspPheValThrArgArgArgIleArgArgThrValValLeuAlaLeu 221
:::: :::: :::: :::: :::: ::::

Db	745	ATCCAAAGACTATACATATTTTAACCCGGAAAGCAATTCAGGTACAGATTTTCGACGATTCATT	804
Oy	222	LeuProCysGlyAlaGluLeuProGlyAlaArgProTyrAlaIleMetValaTyrTyrIleLeuAsp	241
Db	805	CAGCAATTCAGTGAATGTAAAGCCATCGCCAGAAACCTTAACCTTAAGTATCTTATTAAC	864
Oy	242	LeuGluArgLeuHisProAlaAlaThrTyrGluThrPheArgVal	256
Db	865	CTGGAAACCCCTGGAGTTCGCTTCTTAACACGAGAACGTTTGAAGTAAAGAAATCTGCAAGA	924
Oy	257	GlyLeuProGlyAlaGluGluGluLeuProGlyLeuLeuArgValAlaGlyAspAsnGlyIle	276
Db	925	GGT---CCTTCAGGTAGAGGAGATTTTTCGACCATTTAATAAAGCTGGAACGGTGGAAAT	981
Oy	277	ProTrpSerSer-----AsnAspGluLeuPheGlnThr	287
Db	982	CAGTGGTCAAGAGGAAACATAGAAGAAAGTGAACACTGACAGACAGAACGATACAGTTA	104
Oy	288	PheCysAspPheProGlyIleValaIleAspValSerIleAsnGlnAlaProAlaGlyIlePro	307
Db	1042	TATTTGGATTTCCCTGATATTTATGTATGTCAGATTAAGCAAGAAACAG---GAATGC	1091
Oy	308	AlaGlyGluHisArgLeuValThrValThraArgLeuAspGlnHisIleLeuGluAlaGlu	327
Db	1099	TCAAAATGAAGATAGATTTGATTAAGTACATGCTCCATTAACAAAGTGTAAAGTTTGGAGATACGA	1155
Oy	328	PheProGlyIleProGluIleLeuSerPheValAlaLeuValAspGlyTyrPheArgLeu	347
Db	1159	CTTAAGCTCATTTAAAGAAAGCCCTTGATTCATTCGTCATTAATTAAGCCGGATATTACAGACTA	1218
Oy	348	IleCysAspSerArgHisIleTyrPheCysGlyGluValAlaProProAlaGluLeuGluGlu	367
Db	1219	ACTGGGATGGCCACCATTTACCTCGCAAAAGAGTGGCTCCCCAGCTGCTCGAGAAC	1276
Oy	368	GluAlaAspValCysHisGlyProIleThrLeuAspPheAlaIleHisLeuLeuVala	387
Db	1279	ATACACAGCACTGCGCCAGCCCATATACATGATTTTCCATTTGCAAACTAAAGAAAG	1338
Oy	388	AlaGlySerLeuProGlyIleTyrIleIleuArgArgSerProGlnAspTyrAspSerPhe	407
Db	1339	GGGGTAACCAACACTGACACTATATGTGCTCGACGATCCAGCCCTTAAGACCTTCAACAAATAC	1398
Oy	408	LeuLeuThrAlaCysValGlnThrProLeuGlyProAspTyrIleGlyCysLeuIleArg	427
Db	1399	TTTTTGACCTTTGCTGTGGACGCGAAATATGTCATTTGAATTAACACATGTTTGGATTACG	1458
Oy	428	GlnAspProSerGlyAlaPheSerLeuValGlyLeuSerGlnProHisArgSerLeuArg	447
Db	1459	AAGATGTGAATGAGGATATCAACCTCAGGGGAGTAAAGAGAACTTCAGTAACCTTTAAG	1518
Oy	448	GluLeuLeuAlaAlaCysTTPAsnSerGlyLeuArgValaIleValaLeuTyrIleu	467
Db	1519	GACCTTTTGAATTCCTACCGAGATGGAACATGTCGCTCGACAGACAGATCATCTTTCAGGTTT	1578
Oy	468	ThrSerCysAlaProArgProGlyGlyLeuSerAsnLeuIleValValaArgArgIle	487
Db	1579	ACCAATATGCTCCGCCCAAGCCAAAGATTAATCAACACTTCTGTCTTCAGA-----	1632
Oy	488	CysAsnProAlaProAlaProGlyCysSerProSerCysCysAla-----LeuThr	504
Db	1633	ACAATATGATTTTGTATGTTCAGATCTACCAACCATTTACAGAGCCATTAATATGTGAAT	1692
Oy	505	GlnLeuSerPheHisThrIleProThrAspSerLeuGluTyrHisGluAsnLeuGlyHis	524
Db	1693	CAATATGCTTTTCAAAATACAGCAATGAAGATTTTAATTAATTAAGAAAGCTTGGCCAA	1752
Oy	525	GlySerPheThrIleHisPheArgGlySerArgArgGluValaIleAsp---GlyGluThr	543
Db	1753	GGTACTTTTCAAAAAATTTTAAAGGTGAAGAAGAAAGTTGGAGATTAAGTCAACTG	1812
Oy	544	HisAspSerGluValIleuLeuIleuValaIleAspSerArgHisAspAsnCysMetGluSer	563
Db	1813	CACAAACGGAAGTCTTTTGAAGAGCTTAAGTAAGCAATAGCACTATTCAGACTT	1872

QY	564	PheLeuGluAlaAlaIaSerLeuMetSerGlnValSerTyrProHisLeuValIleLeuHis	583
Db	1873	TTCTTCGAACACACAGCATGTATGTAGTCAGCTTTCTTCACAGCATTTGGTTTGAATTTAT	1933
QY	584	GlyAlaLysMetIaGly---AspSerIleMetValGlnGluPheValIlyTyrLeuGlyAla	602
Db	1993	GGCTTCGTGTCTGTGTGAAGAGAGAAACATCTGGTTCAAGAAATTTGMAAATTTGGATCA	1993
QY	603	IleAspMetTyrLeuArgLysArgLysLeuValSerAlaSerTyrPheLysGlnVal	622
Db	1993	CTGGATACATACCTCGAAGAAACAAATAATTCATTAATATTATTTGAAACTTGGAGTC	2055
QY	623	ThrIleGlnLeuAlaTyrIleAlaLeuAsnTyrLeuGlnLysPheLysGlyLeuProHisLysAsn	642
Db	2053	GCTAAGCAGTTGGCATGGGCGCATTTGATTTCTAAGAAAAATCCCTTTATCATGGGAAT	2112
QY	643	ValSerAlaArgLysValIleLeuValIaArgLysGly-----AspGlyAsnPro	660
Db	2113	GTGTGTGGTAAAAATATTCCTGTTATCGACAAACAGACAGAAACGGGCAACCCACT	2172
QY	661	PheIleLysLeuSerAspProGlyAlaSerProThrValLeuSerLeuGlnMetLeuThr	680
Db	2173	TTTATCAAACTTAGATATCCGGCATTTAGCATTTACACTTCTACGAAAGACATTTCTCAG	2232
QY	681	AspArgIleProTyrValAlaProGlnLysLeuGlnIleAlaIleThrLeuLysLeuIleu	700
Db	2233	GAGAGAAATACCATGGGTACCTCTCATGACATTGAGATCTCTAAATATTCATATGGCA	2292
QY	701	AlaAspLysTyrPheGlyPheGlyAlaIleThrTyrPheLysValPheGlnArgGlyProAlaHis	720
Db	2293	ACAGACAACTGGAGCTTGGGAGCACCTCTGGGAGATCTGCAGTGGAGACATTAAGCC	2352
QY	721	IleThrSerLeuGlnProAlaLysLysLeuLysPheTyrGlnAspGlnGlyGlnLeuPro	740
Db	2353	CTGAGTGTCTGTGATTTCTCAAGAAACCTGACAGTTCTATGAAAGATACATCAGCTTCT	2412
QY	741	AlaLeuLysTyrPheThrGlnLeuAlaGlyLeuIleThrGlnLysMetAlaTyrAspProGly	760
Db	2413	GCACCCAACTGACACAGATTTAGCAACCTTATTAATATTCATGGACATATGAGCCAGAT	2472
QY	761	ArgArgProSerPheArgAlaIleLeuAlaArgLeuAsnGlyLeuIleThrSerAspTyr	780
Db	2473	TTTCAGGCGCTTCAGAGCTGTATCCGTGATCTTAAACAGCTCTTTTACTCCAGATTAT	2532
QY	781	GluLeuLeuSerAspProThrProGlyIleProSerProArg---AspGluLeuLysVal	799
Db	2533	GAACTACATTAAGAA---AATGACATGCTACCAACATGACAAATAGTGGCCCTTAGGGTTT	2589
QY	800	AlaGlyIaGlnIleLeuTyrAlaCysGlnAspProAlaIlePheGlnGlnIaGlnHisLeuLys	819
Db	2590	TCGTGGTCT-----TTTGAAGACAGGAGACCCCTACACAGATTGGAAGAGACACTTGAAG	2643
QY	820	TyrIleSerLeuLeuGlyLysGlyAsnPheGlySerValGlyLeuLysArgTyrAspPro	839
Db	2644	TTTCTACAGCGAGCTTTGGCAAAAGTACTTCGGAGTGTGAAGATGTGGCGCTATGACCCG	2703
QY	840	IleuGlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGlnHisSerValProAsp	859
Db	2704	CTGACAGCAACACTGGCGAGGTGTCCGTGTGAAGAAACTCCAGCACACACTGAAGAG	2763
QY	860	GlnGlnArgAspPheGlnArgGlnGlyIleGlnIleLeuLysAlaIleuHisSerAspPheIle	879
Db	2764	CACCCCGAGACTTTAGAGAGAGATCCGAGATCTCGAAATCTCTTGACAGCATCAACATTC	2823
QY	880	ValLysTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuArgLeuValMetGly	899
Db	2824	GTCAAGATCAAGAGAGATGTCTTACAGTCCGGGTGGCGCAACCTAAGATTAATTATGGA	2883
QY	900	TyrLeuProSerGlyLysLysArgAspLeuLeuGlnArgHisArg---GlyLeuHisThr	918
Db	2884	TATTTACCATATGGAAGTTTACGAGACTATCTCCAAAAAACCTTAAGACAGGATAGATCAC	2943

```

OY 919 AspartgLeuLeuPhealatrpginileCysLyseGlyWetGluTyrleuGlyAlaArg 938
    :::::::::: ::::::::::::::::::::::::::::::::::::::::::::
Db 2944 AAAAAATTCTTCAATACACATCTCAGATATGCAGAGGCATGGAATATCTGTGTACAAA 3003
OY 939 ArgCysValHisArgaspLeuAlaAlaArgAsnIleLeuValGluSerGluAlaHisVal 958
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 3004 AGGTATATCTCCACAGGCGCTCGCACACAGACATATATGTTGGAATAATGGAACAGGGTT 3063
OY 959 LysIleLeaspPheGlyLeuAlaLysLeuLeuProleuGlyLysAspTyrTyrValVal 978
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 3064 AAATATAGACAGCTTCGATTAACCAAGTCTGCCGACAGCAAGAATATCTACAAAGTA 3123
OY 979 ArgGluProGlyGlnSerProIlePheTyrAlaProGluSerLeuSerAspAsnIle 988
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 3124 AAGAGCCAGGAGGAAACCCCATATCTGTACCCACCTAATCTTGACGGAGAGCAAG 3183
OY 999 PheSerArgGlnSerAspValTyrPheSerPheGlyValValLeuTyrGluLeuPheThrTyr 1018
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 3184 TTTTCTGGCCCTCAGATGTGTGAGCTTTGAGACTGTTCTATACGAACCTTTCCACATAC 3243
OY 1019 CysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetGlyProGluArgGlu 1038
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 3244 ATCGAGAAGAGTAAAGTCACCCGTGGAATTTATGCGAATGATGTCGAATGATTAACAA 3303
OY 1039 GlyProProLeuGlyCys---ArgLeuLeuGluLeuLeuAlaGluGlyArgArgLeuProPro 1057
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 3304 GGGCAATGATGTGTTCATTTGATGATAGAGCTACTGAAAGACACAGGAAGATGCCAAG 3363
OY 1058 ProProThrCysProThrGluValGlnGluLeuMetGlnLeuCysTTPAlaProGluPro 1077
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 3364 CCAGAGAGATGCCCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3423
OY 1078 HisAspArgProAlaPheAlaThrLeuSer 1087
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 3424 AGCCAGCGTCCCTCCTTCAGGAGACTTTTCG 3453

RESULT 9
AADO3607
ID AADO3607 standard; cDNA: 3629 BP.
XX AADO3607;
AC AADO3607;
DT 19-JUN-2001 (first entry)
XX
DE Murine (Janus kinase 2) Jak2 cDNA.
XX
KW Mouse; Janus kinase 2; Jak2; therapy; cytokine; tyrosine kinase;
KW antiproliferative; cytostatic; cell proliferative disorder;
KW cellular response; erythrocytosis; ss.
XX
OS Mus sp.
XX
FH Key 1.93 Location/Qualifiers
FT 5'UTR /tag= a
FT CDS 94..3483 /tag= b
FT /product= "Murine (Janus kinase 2) Jak2 protein"
FT misc-feature 523
FT /tag= c
FT /note= "5' end of published partial Jak2 cDNA
FT sequence (Harpur et al.)"
FT conflict /tag= d
FT replace (551..553, CCC)
FT /tag= e
FT replace (1089, T)
FT /tag= f
FT replace (1103, C)
FT /tag= g
FT replace (1111, G)
FT /tag= h
FT replace (1119, G)
FT /tag= i
FT replace (1122, C)
FT conflict

FT conflict /tag= i
FT replace (1128, C)
FT /tag= j
FT replace (1131, G)
FT /tag= k
FT replace (1134, G)
FT /tag= l
FT replace (1137, C)
FT /tag= m
FT replace (1140, G)
FT /tag= n
FT replace (1143, G)
FT /tag= o
FT replace (1146, C)
FT /tag= p
FT replace (1188, T)
FT /tag= q
FT replace (1194, G)
FT /tag= r
FT replace (1230, G)
FT /tag= s
FT replace (1245, T)
FT /tag= t
FT replace (1260, T)
FT /tag= u
FT replace (1266, C)
FT /tag= v
FT replace (1272, T)
FT /tag= w
FT replace (1275, Q)
FT /tag= x
FT replace (1293, T)
FT /tag= y
FT replace (1305, T)
FT /tag= z
FT replace (1323, C)
FT /tag= aa
FT replace (1341, A)
FT /tag= ab
FT replace (1344, A)
FT /tag= ac
FT replace (1359, G)
FT /tag= ad
FT replace (1365, Q)
FT /tag= ae
FT replace (1368, T)
FT /tag= af
FT replace (1374, T)
FT /tag= ag
FT replace (1401, C)
FT /tag= ah
FT replace (1413, C)
FT /tag= ai
FT replace (1431, T)
FT /tag= aj
FT replace (1453, Q)
FT /tag= ak
FT replace (1476, G)
FT /tag= al
FT replace (1488, T)
FT /tag= am
FT replace (1511..1512, GT)
FT /tag= an
FT replace (1578, C)
FT /tag= ao
FT replace (1590, T)
FT /tag= ap
FT replace (1593, T)
FT /tag= aq
FT replace (1602, G)
FT /tag= ar
FT replace (1623, T)
FT /tag= as

```

```

FT conflict replace (1642, G)
FT /*tag- at
FT replace (1657, C)
FT /*tag- au
FT replace (1728, G)
FT /*tag- av
FT replace (1743, C)
FT /*tag- aw
FT replace (1755, C)
FT /*tag- ax
FT replace (1770, A)
FT /*tag- ay
FT replace (1809, G)
FT /*tag- az
FT replace (1816, G)
FT /*tag- ba
FT replace (1821, C)
FT /*tag- bb
FT replace (1857, A)
FT /*tag- bc
FT replace (1878, T)
FT /*tag- bd
FT replace (1935, A)
FT /*tag- be
FT replace (1938, A)
FT /*tag- bf
FT replace (1963, T)
FT /*tag- bg
FT replace (1974, G)
FT /*tag- bh
FT replace (2025, T)
FT /*tag- bi
FT replace (2055, G)
FT /*tag- bj
FT replace (2079, C)
FT /*tag- bk
FT replace (2082, C)
FT /*tag- bl
FT misc_feature
FT 2226
FT /*tag- bm
FT /note- "Nucleotides corresponding to 7 amino acid insert"
FT replace (2253, A)
FT /*tag- bn
FT replace (2259, G)
FT /*tag- bo
FT replace (2283, A)
FT /*tag- bp
FT replace (2285, C)
FT /*tag- bq
FT replace (2433, G)
FT /*tag- br
FT replace (3453, C)
FT /*tag- bs
FT 3484.3629
FT /*tag- bt
FT replace (3579, C)
FT /*tag- bu
FT replace (3595, A)
FT /*tag- bv
FT replace (3599, A)
FT /*tag- bw
FT replace (3620, A)
FT /*tag- bx
FT conflict
FT US6210654-B1.
FT 03-APR-2001.
FT 08-OCT-1997;
FT 18-JUN-1996;
FT 29-JUL-1993;
FT 96US-0665574.
FT 93US-0097997.

```

```

PA (SUDD-) ST JUDE CHILDREN'S HOSPITAL.
XX
XX Ihle J, Witthuhn BA, Quelle FW, Silvennoinen O;
PI WPI; 2001-265367/27.
XX
XX P-PSDB; AAE00352.
XX
XX Modulating a biological response mediated by Jak kinase 2 activation to
PT a cytokine, useful for treating excessive proliferation of eukaryotic
PT cells, comprises inhibiting or enhancing tyrosine kinase activity of
PT Jak kinase in the cell.
XX
XX Example 1: Fig 1, 100pp; English.
XX
XX The present sequence is a cDNA encoding murine (Janus kinase 2) Jak2
CC tyrosine kinase. Jak2 sequence has a 600 amino acid long N-terminus that
CC lacks obvious SH2 (src homology 2) and SH3 domains. Following this is a
CC kinase related domain (domain 2) and a carboxyl kinase domain (domain
CC 1). Jak kinases mediate cytokine activity through their tyrosine
Alignment Scores:
Pred. No.: 2,33e-190 Length: 3629
Score: 2619.50 Matches: 537
Percent Similarity: 65.50% Conservative: 190
Best Local Similarity: 48.38% Mismatches: 350
Query Match: 44.70% Indels: 33
DB: 22 Gaps: 15
US-09-397-967-16 (1-1099) x AAD03607 (1-3629)
QY 4 ProserGluGluThrProLeuIleProGlnArgSerCysSerLeuSerSerGluAla 23
DB 145 CCTGATACATCAGATGGATGATATTCCTGGAATGCTAATTCGTAGACAGATATAGCCA 204
QY 24 GlYAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAl 43
DB 205 GTCTCTCAAGTATGATCTGTACCATTCCTGTGGCAAGCTGAGAGAGATGTGAGATT 264
QY 44 SerPheGlyAspTyrLeuAlaGluAspLeuSerCysValArgAlaAlaLysAlaCysGlyIle 63
DB 265 CCAAGTGGAGAGATATGCTTCACAGAAATTTGTGTGGCTGCTTCAAACTTGCTGTATT 324
QY 64 LeuProValTyrHisSerLeuPheAlaLeuAlaIleAlaIleAlaIleAlaIleAlaIle 83
DB 325 ACGCTGTATATCAATATATATGTTGCGTAAATGAGGAAACCGAAGATGCTGATACCA 384
QY 84 ProSerHisIlePheCysIleGluAspValAspThrGlnValAlaLeuValTyrArgLeuArg 103
DB 385 CCCAATCATGCTCTCCACATAGACAGTCAMCCAGCATGACATGATCTACAGATTAAGG 444
QY 104 PheTyrPheProAspTyrPhe-----GlyLeuGluIleThrCysHisArgPheGlyLeuArg 121
DB 445 TTCTACTTCCCTCATTTGGTACTGTAGTGGACAGACAGAACCTACAGATAGCGAGTCC 504
QY 122 LysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLeuPheAlaGlnHis 141
DB 505 CTTGGGGCTGAGAGCTGCTGCTGATGACTTTGCATGCTGTACTTTTCTGACAGTGG 564
QY 142 ArgSerAspLeuValSerGlyArgLeuProValIleGlyLeuSerMetLysGluGlnGly 161
DB 565 CGGCATGATTTGTTACGAGATGATAAAGTACCTGTACATCAATCAAGTGAAGAG 624
QY 162 PheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGlnAlaGlnArgPro 181
DB 625 TGTCTTGGATGGCGGTGTAGACATGATGAGATATAGCTAAGAGAAACAGCACTCA 684
QY 182 GlyGluLeuLeuLysThrValSerTyrLysAlaCysLeuProProSerLeuArgAspVal 201
DB 685 CTGCGTGTATTAACCTGTACAGTACAAAGACTTCTTACCAAAAGCTGTCGACCGAAG 744
QY 202 IleGlnGlyGlnAsnPheValThrArgArgArgIleArgArgThrValValLeuAlaLeu 221
DB 745 ATCCAAGACTATACATTTTAACCCGGAAGCGAATCAGGTACAGATTTCGAGATTTCATT 804

```


Dh	655	GCCCTGTCATTCGTGCATTAATTACGGGATTTACAGACTAACTGGCGATGGCGACCAT	714
Oy	354	TYrPheCysLysGluValAlaProGlarGleuLeuGluGluValAlaAspValCysHis	373
Dh	715	TACCTCGCAAAAGAGGTGGCTCCCGACGTGGCTCGGAACATACACACCACTCCGAC	774
Oy	374	GLYProIleThrLeuAspPheAlaIleHisLysLeuLysAlaAlaGlySerLeuProGly	393
Dh	775	GGCCCATATTCATGATGATTTTCCATTGCAACAACTAAAGAAAGGGGGGTAAACAGACTGG	834
Oy	394	ThrTrIleLeuArgArgSerProGlnAspTyrAspSerPheLeuLeuThrAlaCysVal	413
Dh	835	CTATATGTGTACGATGACGCCCTTAAGACTTCACAAATATCTTGTGACCTTGGCTGTT	894
Oy	414	GLnThrProLeuGlyProAspTyrLysGlyLysLeuIleArgGlnAspProSerGlyAla	433
Dh	895	GAGCGAAGAAATGCTCATGTAAATAAACCTGTGTGATTTGCAAGAAATGAGATGAGAA	954
Oy	434	PheSerLeuValGlyLeuSerGlnProHisArgSerLeuArgGluLeuLeuAlaIaCys	453
Dh	955	TACACCTCAGCGGACTATAGAACTTCAGTAACTTAAGAGACTTTGAATGTGCTAC	1014
Oy	454	TrpAsnSerGlyLeuArgValAspGlyAlaIleLeuTyrLeuThrSerCysGlyAlaPro	473
Dh	1015	CAGATGGAAACTGTGGCGCTCAGCACTATCATCTTCCAGTTTACCAATGCTGCCCCCA	1074
Oy	474	ArgProLysGluLysSerAsnLeuIleValAlaArgGlyCysAsnProAlaProAla	493
Dh	1075	AAGCCAAAGATTAATCAAACTCTTCGCTTCACA-----ACAAATGGATTTGTGAT	1128
Oy	494	ProGlyCysSerProSerCysCysAla-----LeuThrGlnLeuSerPheHisThr	510
Dh	1129	GTTGAGATCTCCACCATTCATCAGAGCGCATTAATTAATGATCAAACTGCTTTCCACAA	1188
Oy	511	IleProThrAspSerLeuGluIleTrpHisGluAsnLeuGluIleGlySerPheThrLysIle	530
Dh	1189	ATCAGGAATGAGATTTAATATTATTAAGAAGCTGTGGCCAAAGTACTTTACAAAATTT	1248
Oy	531	PheArgGlySerIleArgGluValAlaAsp--GlyIleTrpHisAspSerGluValLeu	549
Dh	1249	TTTAAAGGTGAAGAAGAAAGTTGAGATATTATGCTCAACTCACAACAAACGGAAGTCTT	1308
Oy	550	LeuLysValIleLysSerAspGlnHisAspArgAsnCysMetGluSerPheLeuGlnAlaAlaSer	569
Dh	1309	TTGAAAGCTCCTGAATAAAGCAATACAGAACTATTAAGAAGCTTTCTTTCGAAGCGACAGC	1368
Oy	570	LeuMetSerGlnValSerTyrProHisLeuValLeuLeuHisGlyValCysMetAlaGly	589
Dh	1369	ATGATAGACTCAGCTTCTTCCAAAGCAATTTGGTTGAATTAATGAGTGTGTGTCTGTGG	1428
Oy	590	---AspSerIleMetValGlnGluProPheValTyrLeuGlyAlaIleAspMetTyrLeuArg	608
Dh	1429	GAGGAGAACATTTCTGGTTCAAGAAATTTGTAAATTTGGATGACAGTCAATCACTGAG	1488
Oy	609	LysArgGlyHisLeuValSerAlaSerTyrPheLysGluValThrLysGlnLeuAlaTyr	628
Dh	1489	AAGAACAAAATTCCTCATTAATATTTATGGAACCTTGGAGTGGGTAAAGAGTGGCAATCG	1548
Oy	629	AlaLeuAsnTyrLeuGluAspLysGlyLeuProHisGlyAsnValSerAlaArgLysAla	648
Dh	1549	GCCATGCAATTTCTTAGAAGAAATTCCTTATTCATGAGGAATGTGTGTCAAAAATATC	1608
Oy	649	LeuLeuAlaArgGluGlyGly-----AspGlyAsnProProPheIleLysLeuSerAsp	666
Dh	1609	CTGCTTATCAGGAAGAAAGACAGGAGACGGGAGACCCACTTTCATCAAACTTGAT	1668
Oy	667	ProGlyValSerProThrValLeuSerLeuGluMetLeuThrAspArgIleProThrVal	686
Dh	1669	CCTGGCATTTAGCATTAACACTTCACGAAGACATCTTCTAGAGAGATATACATGAGCTA	1728
Oy	687	AlaProGluCysLeuGlnGluAlaGlnThrLeuCysLeuGluAlaAspLysTyrPheLys	706
Dh	1729	CTCTGATGATGATTTGAAGATCTCAAAATCTCAATCTGGCAACACAGATGGAGCTTC	1788

QY	707	GIATATrThrTrpGluValPheGlnArgIleProAlaHisIleThrSerLeuGluPro	726
Db	1789	GGGACCACTCTGTGGAGATCTGCAGTGGAGGAGATAAAGCCCTTCAAGTCTTGATTC	1848
QY	727	AlAluSlyLeuLeuSlyPheTrpGluAspGlnIleuGlnLeuProAlaLeuIstPheTrpGlu	746
Db	1849	CAAAAGAACTCAGTCTATGAGATGAATGAAGCATCAAGCTTCTCGCAACCCAAAGTGGACAG	1908
QY	747	LeuAlaGluLeuIleThrGlnCysMetAlaTrpAspProGluIleArgArgProSerPheArg	766
Db	1909	TTAGCAAACTTATAAATTAATGATGACATGACACTATGAGCCGATTTCAAGGCTCTTCAGA	1968
QY	767	AlAlaIleuArgAspLeuAsnGlnLeuIleThrSerAspTrpGluLeuLeuSerAspPro	786
Db	1969	GGTGTATCCCTCGATCTTTAAAGACCTCTTTACTCCAGATTATGAAGACTACTAACAGAA	2025
QY	787	ThrProGluIleProSerProArg---AspGluLeuCysValAlaGluIleuLeuTr	805
Db	2026	AATGATATCTTACCAAAACATGAGAATAGTGGCCCTGGGTTTCTGGTGGCT---TTT	2079
QY	806	AlAcSGlnAspProAlaIlePheGlnIleuArgIleLeuSlyLeuIstPheSerLeuLeuGly	825
Db	2080	GAAGACAGGAGCCCTACACAGTTTGAAGAGAGACACTTGAAGTTTACAGACAGCTTGGC	2139
QY	826	LysGluAsnPheGlySerValGluLeuCysArgTrpAspProLeuGluIAspAsnThrGly	845
Db	2140	AAAGTAAGTCTGGGAGTGGAGATGTGCCGTATGACCCGCTGAGGACAAACACTGGC	2199
QY	846	ProLeuValAlaValAluSglnLeuGlnHisSerValProAspGlnIleuArgAspPheGln	865
Db	2200	GAGTGTGTGCTGTGAAGAAATCCACAGCAAGCACTGAAGAACACTCCGAGACTTTGAG	2259
QY	866	ArgGluIleGlnIleLeuSlyAsnAlaLeuHisSerAspPheIleValIstPheArgGluVal	885
Db	2260	AGGAGATCTCAGATCTCTGAATCTCTTGCAGCATGACAACTGCTCAAGTAAAGGGAGT	2319
QY	886	SerTrpClyProGluIleArgIleSerLeuValGluValMetGluTrpLeuProSerGlyCys	905
Db	2320	TGCTACAGTGGGGTGGCGGCCCACTTAAGATTAATTAATGAAATTTATCCATGTGAAGT	2379
QY	906	LeuArgAspLeuLeuGlnArgHisArg---GlyLeuHisThrAspArgLeuLeuLeuPhe	924
Db	2380	TTTACGAACTATCTCCAAAACATTAAGAAAGACGATAGATACAAAAAATCTTCTCAATAC	2439
QY	925	AlATrPglIleCysIstGlyMetGluTrpLeuGluValAlaArgArgCysValHisArgAsp	944
Db	2440	ACATTCAGATATGCAAGGGCATGGAATATCTTGGTCAAAAAGGATATATCCACAGGAC	2499
QY	945	LeuAlaAlaArgAsnIleLeuValGluSerGluAlaHisValIstIleAlaAspPheGly	964
Db	2500	CTGGCAACAAGAAACATATTTGGTGAATAATGACAAACGGGTAAATAGACAGCTTCGCA	2559
QY	965	LeuAlaIstLeuLeuProLeuGluIstAspTrpTrpValValAlaArgGluProGlyGlnSer	984
Db	2560	TTAACCAAACTCTTGGCGGACAGCAAAAGAAATACTACAAAGTAAAGGACGACAGGGAAGC	2619
QY	985	ProIlePheTrpTrpAlaProGluSerLeuSerAspAsnIlePheSerArgGlnSerAsp	1004
Db	2620	CCCATTTCTGGTGGACGACCTCAATCCCTTACGAGAGCAAGTTTCTGTGGCTTCGAT	2679
QY	1005	ValTrpSerPheGlyValAlaLeuTrpGluLeuPheThrTrpCysAspIstSerCysSer	1024
Db	2680	GTGGGAGCTTTGAGTGGTCTTATACCAACTTTTCCATATACATCAGGAAGAGTAAAGT	2739
QY	1025	ProSerAlaGluPheLeuArgMetMetClyProGluArgGluGlyProProLeuCys---	1043
Db	2740	CCACCCGGGAAATTATTCGCGAATGATTTGGCAATGATTAACAAAGGGCAATGATTTGTTC	2799
QY	1044	ArgLeuLeuGluLeuLeuAlaGluGlyArgArgLeuProArgProArgProTrpCysProThr	1063
Db	2800	CATTGTATGAGACTACTGAAGACGACGGAAGATTGGCCCAAGCCGAGAGATGCCCCAGAT	2859

QY 414 GlnThrProLeuGlyProAspTyrLysGlyCysLeuLeuLeuArgAspLeuAspProSerGlyAla 433
 Db 895 GACGAGAAATGTTATTTGAATATAACACTGTTGATTCACAAAGATGAGATGGAGAG 954
 QY 434 PheSerLeuValGlyLeuSerGlnProHisArgSerLeuArgGluLeuLeuAlaAlaCys 453
 Db 955 TACACCTCAGTGGAGCTAGAGAACCTTCAGTACTTCTTAAGAGACCTTTTGAATGGCTAC 1014
 QY 454 TrpAsnSerGlyLeuArgValAspGlyAlaAlaLeuTyrLeuThrSerCysAlaPro 473
 Db 1015 CAGATGGAAACTGCGCTCAGACAGATATCTTCAGTTCACCAATCTGCTCCCA 1074
 QY 474 ArgProLysGluLysSerAsnLeuLeuValArgArgGlyCysAsnProAlaAla 493
 Db 1075 AACCCAAAGATTAATCAAACTCTTGTCTCAGA-----ACAATAGTGTTCGAT 1128
 QY 494 ProGlyCysSerProSerCysAla-----LeuThrGlnLeuSerPheHisThr 510
 Db 1129 GTTCACCTCTCACCACTTACAGAGCATATATGTGATCAATAGGTGTTCACAAA 1188
 QY 511 IleProThrAspSerLeuGluTrpHisGluAsnLeuGlyHisGlySerPheThrLysIle 530
 Db 1189 ATCAGAAATGAAATTTGATATTAAATGAAGCCTTGCCCAAGGCACTTTACAAAATA 1248
 QY 531 PheArgGlySerArgArgGluValAlaAsp--GlyGluThrHisAspSerGluValLeu 549
 Db 1249 TTTAAAGGTAGTAGAGAGAGAGAGTTGAGATTTAGTCAGCTGCACGAAACCGAAGTCTT 1308
 QY 550 LeuLysValMetAspSerArgHisArgAsnGlyMetGluSerPheLeuGluAlaAlaSer 569
 Db 1309 TTGAAAGTCTGATATGAACCATAGAACATTAATTCAGAGCTTCTTGAAGAGCAAGC 1368
 QY 570 LeuMetSerGlnValSerTyrProHisLeuValLeuHisGlyAlaCysMetAlaGly 589
 Db 1369 ATGATGAGCAGCTTCTCCACAGCATTTGGTTGATTAATGAGATGTGCTGTGGA 1428
 QY 590 ---AspSerIleMetValGlnGluPheValTyrLeuGlyAlaAlaLeuAspMetTyrLeuArg 608
 Db 1429 GAGGAAACATTTTGGTTCAGAGTTGTAATAATTTGGATCACTGATACATCCGGAAG 1488
 QY 609 LysArgGlyHisLeuValSerAlaSerTyrLysLeuGlnValThrLysGlnLeuAlaTyr 628
 Db 1489 AAGCAAAAATTTCTTAATATATATATGGAACCTTGAGCTGCCAAGCAGTTGGCATGG 1548
 QY 629 AlaLeuAsnTyrLeuGluAspLysGlyLeuProHisGlyAsnValSerAlaArgLysVal 648
 Db 1549 GCCATGCACTCTCCGAAAGAAAATCCCTTATCATGGAGATGTGTGCTAAATAATATC 1608
 QY 649 LeuLeuAlaArgGluGly-----AspGlyAsnProProPheLeuLysLeuSerAsp 666
 Db 1609 CTGCTTATCAGAGAAAGAGACAGAGACGGGAGACCCACTTTCATCAAACTTAATGAT 1668
 QY 667 ProGlyValSerProThrValLeu-----SerLeuGluMetLeu 679
 Db 1669 CCTGCATTAGCATTTACACTTTCACCGAAGACATTTCTCCGTGTTCCTCCAGTTCTT 1728
 QY 680 ThrAspArgIleProTrpValAlaProGluCysLeuGlnGluAlaGlnThrLeuCysLeu 699
 Db 1729 CAGGAGAGAAATACATAGGTGTACACCTGAGTGATGAGATCTTAATAATTAACCTG 1788
 QY 700 GluAlaAspLysTyrGlyPheGlyAlaThrThrTrpGluValPheGlnArgGlyProAla 719
 Db 1789 GCAACAGCAAGATGAGCTTGGGACCACTGTGGAGATCTGCGAGATGCGAGAGAGATAAG 1848
 QY 720 HisIleThrSerLeuGluProAlaLysLysLeuLysPheThrGluAspGlnGlyGlnLeu 739
 Db 1849 CCCCTGAGTGTCTGATTTCTCAAGAAAGCTGAGTTCTTAAGAAATAGATGACGCTT 1908
 QY 740 ProAlaLeuLysTrpThrGluLeuAlaGlyLeuIleThrGlnCysMetAlaTyrAspPro 759
 Db 1909 CCTGACCCCAAGTGGAGACAGAGTTGGCAACCTTATTAATTAATTCATGAGCATATGAGCCA 1968

QY 760 GlyArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIleThrSerAsp 779
 Db 1969 GATTTCAGGCTCTGTTCAAGAGCTGATCATCCGATCTTACAGCCGTTACTCCGAT 2028
 QY 780 TyrGluLeuLeuSerAspProThrProGlyIleProSerProArg--AspGluLeuCys 798
 Db 2029 TATGAACTACTACAGAA-----AATGACATGCTACCAAAACATGAAATAGCTGCTCAGG 2085
 QY 799 ValAlaGlyValGlnLeuTyrAlaCysGlnAspProAlaIlePheGluGluArgHisLeu 818
 Db 2086 TTTTCTGTCT-----TTGAAAGACAGGAGCCCTACACAGTTCAGAGAGACACTTG 2139
 QY 819 LysTyrIleSerLeuLeuGlyLysGlyAsnPheGlySerValGluLeuCysArgTyrAsp 838
 Db 2140 AAGTTCTACAGCAGCTTGGCAAAAGTAACCTGGGAGTGTGGAGATGTCGCCCTATGAC 2199
 QY 839 ProLeuGlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGlnHisSerValPro 858
 Db 2200 CCGCTGAGAGCAAACTGGCGAGGTGTCTGTGTGAAGAACTCCAGCACACACTGAA 2259
 QY 859 AspGlnGlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuHisSerAspPhe 878
 Db 2260 GACACCTCCGAGACTTTGAGAGGAGATCGAGATCTGAATCTTCAGCATGACAAAC 2319
 QY 879 IleValLysTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuArgLeuValMet 898
 Db 2320 ATCGTCAGATACAGAGGAGGTGCTACAGTGGCGGTGGCGCAACCTAGATTAATATG 2379
 QY 899 GluTyrLeuProSerGlyCysLeuArgAspLeuLeuGlnArgHisArg--GlyLeuHis 917
 Db 2380 GAATATTATTCATATGGAAGTTCAGAGACTATCTCCAAAACATGAAGACGAGATAGAT 2439
 QY 918 ThrAspArgLeuLeuLeuPheAlaTrpGlnIleCysLysGlyMetGluTyrLeuGlyAla 937
 Db 2440 CACAAAATCTTTCATATACATCATCTCAGATATGCAAGGCAATGGAATATCTTGATACA 2499
 QY 938 ArgArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGluSerGluAlaHis 957
 Db 2500 AAAAGTATATCCACAGGCGCCTGCAACAGCAACATATTTGCTGGAATAATGAAACAG 2559
 QY 958 ValLysIleLeuAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAspTyrTyrVal 977
 Db 2560 GTTAAATAGAGACTTTCGATTAACCAAGTCTTGCCGAGAGCAAAAGATATCTACAAA 2619
 QY 978 ValArgGluProGlyLysLeuProIlePheTrpTyrAlaProGluSerLeuSerAspAsn 997
 Db 2620 GTAAGGAGCGAGGAGAAAGCCCATATTTCTGTAGCGACCTGAATCCTTGACGAGAG 2679
 QY 998 IlePheSerArgGlnSerAspValTrpSerPheGlyValValLeuTyrGluLeuPheThr 1017
 Db 2680 AAGTTTCTGTGGCTTCAGATGTGTGAGCTTGGAGTGTCTTATACGAACCTTTTCACA 2739
 QY 1018 TyrCysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetGlyProGluArg 1037
 Db 2740 TACATGAGAGAGAGTAAAGTCCACCCGAGAAATTTATGCGAATGATTTGGCAATGATAAA 2799
 QY 1057 ProProProThrCysProThrGluValGlnGluLeuMetGlnLeuCysTrpAlaProGlu 1076
 Db 2860 AGGCCAGAGAGAGCCCAATGAGATTTATGTGATATGACAGAGTGTGGAACAACAAT 2919
 QY 1077 ProHisAspArgProAlaPheAlaThrLeuSer 1087
 Db 2920 GTGAGCAGCGTCCCTTCACGAGACTTTCC 2952
 RESULT 12
 ABR72331
 ID ABR72331 standard; DNA; 4191 BP.
 XX
 AC ABR72331;

XX 30-JUL-2002 (first entry)
 XX DNA encoding lymphoma associated mouse JAK1 protein.
 DE
 XX Lymphoma associated protein: PIK3R1, GNMS; XI-(alpha-s-); NESP55; JAK1;
 KW Neurogranin; Nf12; Lymphoma; DNA vaccine; cytotoxic T-cell;
 KW animal model; Leukemia; ds.
 XX Mus sp.
 OS
 XX WO200224867-A2.
 PN
 XX 28-MAR-2002.
 PD
 XX 24-SEP-2001; 2001US-0529798.
 PE
 XX 22-SEP-2000; 2000US-0668644.
 PR 13-JUL-2001; 2001US-0905390.
 PR 13-JUL-2001; 2001US-0905491.
 XX
 PA (UYAA-) UNIV AARHUS.
 XX
 PT Pedersen FS, Sorensen AB, Hernandez JM, Nielsen AA, Moring HO;
 DR WPI: 2002-416279/44.
 DR P-PSDB: ABG60307.
 XX
 XX Novel recombinant lymphoma associated protein (LAP) such as PIK3R1,
 PT GNMS, JAK1, Neurogranin, Nf12 proteins, useful for identifying
 PT inhibitors of LAP activity that are used for treating lymphoma -
 XX
 PS Claim 1: Page 148; 160pp; English.
 XX
 CC The invention describes a recombinant protein (I) from a lymphoma
 CC associated protein (LAP) sequence of PIK3R1, GNMS (including
 CC XI-(alpha-s-), and NESP55), JAK1, Neurogranin or Nf12 proteins. (I) is
 CC useful for screening for a bioactive agent capable of binding to an LA
 CC protein (LAP) which is encoded by a polynucleotide (II) and (II) is
 CC useful for evaluating the effect of a candidate lymphoma drug in a
 CC patient. (I) and (II) is also useful for diagnosing lymphoma involving
 CC determining the expression of one or more (II), or (I) encoded by (II)
 CC in a first tissue type of a first individual and comparing the expression
 CC of the gene(s) from a second normal tissue type from the first individual
 CC or a second unaffected individual, where a difference in the expression
 CC indicates that the first individual has lymphoma. (II) is useful in the
 CC diagnosis and treatment of lymphoma and leukaemia. A host cell
 CC comprising (II) is useful for screening drug candidates which involves
 CC providing the host cell that expresses LA gene, adding a drug candidate
 CC to the cell, and determining the effect of the drug candidate on the
 CC expression of the LA gene. LA genes are also administered as DNA
 CC vaccine, such that expression of the polypeptide encoded by the DNA
 CC vaccine, cytotoxic T-cells and antibodies are induced which recognise and
 CC destroy or eliminate cells expressing LA proteins. The genes are also for
 CC generating animal models of lymphoma. This sequence represents a lymphoma
 CC associated polynucleotide described in the invention.
 XX
 SO Sequence 4191 BP; 1214 A; 982 C; 1024 G; 971 T; 0 other;

Alignment Scores:

Pred. No.: 6,88e-136 Length: 4191
 Score: 1908.50 Matches: 437
 Percent Similarity: 54.52% Conservative: 184
 Best Local Similarity: 38.37% Mismatches: 383
 Query Match: 32.57% Indels: 135
 DB: 24 Gaps: 22

US-09-397-967-16 (1-1099) x ABK72331 (1-4191)

QY 41 LeuserPheSerPheGlyAspTyrLeuAlaGluAspLeuCyValArgAlaIalaIysala 60
 DB 408 CTCGCCCTGGGCGAGAGATATACAGCGGAGAGCTGTGATCATCAGCGCCGCGAGAG 467

QY 61 CysGlyIleLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSerCys 80
 DB 468 TGCAGTATCTCTCTCTGTGTCACAACCTTCGCGCTGTAGCATGACCAAGCTC 527
 QY 81 TrpPheProPheSerHisIlePheCysIleGluAspValAspThrGlnValLeuValTyr 100
 DB 528 TGGTACGGCTCCCAACCGAATCATCACTGTGGATGACAAACGCTCTCCGCTCAGTAC 587
 QY 101 ArgLeuArgPheThrPheProAspTyrPheGlyLeu-----GluThrCysHis 116
 DB 588 CGCATGAGGTTCTACTTTCACCACTGGCAGCAACCAATGACACGAAACAGTGTGATG 647
 QY 117 ArgPheGlyLeuArgLysAspLeu-----ThrSer 126
 DB 648 CGACATTTCTCCAAAGACAGCAAAAACGGCTGTGAGAGAAAGGTTCCAGACCAAC 707
 QY 127 AlaIleLeuAspLeuHisValLeuGluHisLeuPheAlaGlnHisArgSerIleVal 146
 DB 708 CCACCTCTGTGATGCCAGTTCACCTGAGTATCTGTGGACAGGACAGTATGATTTGATC 767
 QY 147 SerGlyArgLeuProValGlyLeuSerMetLysGluGln-----Gly 160
 DB 768 AATGCTCTGGCTCCCATTTGGGAGCCCAAGACGAGCAAGACGACATGATTTGAAAT 827
 QY 161 GluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGlnAlaIleArg 180
 DB 828 GAGTGGCTGGGATGGCGGCTCTCCATCTCCACATGATGATGAAGAATGACG 887
 QY 181 ProGlyGluLeuLeuLysThrValSerTyrLysAlaCysLeuProPheSerLeuArgAsp 200
 DB 888 TTGCGGAACTTCCCAACACATCACTCAAGCATATATTCACAGAAACATTGAATAA 947
 QY 201 ValIleGlnGlyGlnAsnPheValThrArgArgArgIle----- 213
 DB 948 TCCATTCAGACAGAGACCTTCTTACAGAGATGCGAATAATATGTTTTCAGAGATTTC 1007
 QY 214 -----ArgArgThrValValLeuAlaLeuLeuProCysGlyArgLeuPro 228
 DB 1008 TTGAAGAAATTTAACAAACAGACATC-----TGTGAC 1040
 QY 229 GlyArgProTyrAlaLeuMetAlaLysTyrIleLeuAspLeu-----ArgLeuHis 246
 DB 1041 ACGAGTGTCAATGACCTGAGGTGAATACCTGGCTACCTTGGAAACCTTCAATTGACA 1100
 QY 247 ProAlaIleThrThrGluThrPheArgValGlyLeuProGlyAlaGlnGluLeuProGly 266
 DB 1101 AAACATTATGAGCTGAATATTTGTGAGACTTGTATCTGATTTACATCAGAAATGAA 1160
 QY 267 LeuLeuArg-----ValAlaGly 272
 DB 1161 TTGAGTGTGATGCTCATTCGATGACAGTGCATGTTCTGTATGAGGTGATGATGAG 1220
 QY 273 AspAsnGlyIleProIlePheSerAsnAsp----- 282
 DB 1221 AATCTCGGATCCAGTGGCGGAGAAACCAATGTTTCTCTGTGAAAGGAAAAAAT 1280
 QY 282 ----- 282
 DB 1281 AAACGTAGCGGAAAAAAGTGAATATTAACACAGAGATGATGAGAAACAA 1340
 QY 283 -----GluLeuPheGlnThrPheCysAspPheProGluIleValAspValSerIleAsn 300
 DB 1341 CTCGGGAGAGTGAACAAATTTTCTTTCCTGTAATCCACCACTGTAATGAAG 1400
 QY 301 GlnAlaProArgValGlyProAlaGlyGlnHisArgLeuValThrValThrArgMetAsp 320
 DB 1401 GAGTCT-----GTGTCAGCACTTAACAAACAGGAC 1430
 QY 321 GlyHisIleLeuGluAlaGluPheProGlyLeuProGluAlaLeuSerPheValAlaLeu 340
 DB 1431 AACAAACATGAACTCAAGCTCTTCTCGAGAGAAAGCTTGTCTTTGTGCTCCG 1490
 QY 341 ValAspGlyTyrPheArgLeuIleCysAspSerArgHisTyrPheCysLysGluValAla 360

Db 1491 GTGGATGGCTACTTCGCGCTCAGCGAGATGCCACCATTACTCTGTAAGTATGTGGCT 1550

QY 361 ProProArgLeuLeuGluGluGluAlaAspValCysHisGlyProIleThrLeuAspHe 380

Db 1551 CCCCACGACTGTTGCCAATAATACAGAAACGGCTGCCACGGTCAATCTGCACACAAATAT 1610

QY 381 AlaIleHisLysLysLysAlaIaIaGlySerLeuProGlyIleThrLysIleuAlaArgSer 400

Db 1611 GCCATTCATTAAGTGGCGGACAGAAAGGAGGAAGAGGAGATGTAGCTGTGAGTGGAGC 1670

QY 401 ProGlnAspTyrAspSerPheLeuThr---AlaCysValGlnThrPro-----Leu 417

Db 1671 TGCACCGACTTTGACACAACTTCTTATGACGTGCACCTGCTTTGAAAGATCGAGGTATG 1730

QY 418 Gly-----ProAspTyrLysGlyCysLeuIleArgGlnAspProSerGlyValAAspHe 435

Db 1731 GGTGCCCGAAGACGATTCAAGAACTTTACAGATT---GAGGTACAAAGGGCGGTACAGC 1787

QY 436 LeuValGlyLeuSerGlnProHisArgSerLeuArgIleGluIleuAlaIaCysTyrAsn 455

Db 1788 CTGCATGGCTGTATGACCAACTTTCCAGCGCTCGAGACCTCATATGACACCACTCAAGAG 1847

QY 456 SerGlyLeuArgValAspGlyAlaIaLeuTyrLeuThrSerCysCysAlaProArgPro 475

Db 1848 CAGATCCCTGGCCAGGACGACAAACATAGCTTTGTCTTGAACAGATGCTGTACAGCTAAGCT 1907

QY 476 LysGluLysSerAsnLeuIleValValArgArgGlyCysAsnProAlaProAlaProGly 495

Db 1908 CGAGAAATCTCAATCTGCTCGTACGCCACTAAG-----AAAGCCGACAG 1952

QY 496 CysSerProSerCysAlaLeuThrGlnLeuSerPheHisThrIleProThrAspSer 515

Db 1953 TGGCAGCT---GTCTACTCTCATGAGCCAGCACTGATGGATGGAGCTTAAGAAACAT 2009

QY 516 LeuGluTyrHisGluAsnLeuGlnLysGlySerPheThrLysIlePheArgGlySerArg 535

Db 2010 ATTATACAAAGGTAGACACTTGGCAGAGGACACAAGAACATATCTATTCTGGACCTG 2069

QY 536 ArgGluValValAsp-----GlyGluThrHisAspSerCysValLeuLeuLys 551

Db 2070 CTGGACTACAGAGATGAGGAAGAAATTTGCTGAAGAGAAAGATAAAGTATCTCTCAA 2129

QY 552 ValMetAspSerArgHisAspAsnGlyMetGluSerPheLeuGlnAlaIaSerLeuMet 571

Db 2130 GTCTTAGACCCCAACCGCGGACATCTCTGTGCGCTTTTGAAGGCTGATGAGTATG 2189

QY 572 SerGlnValSerTyrProHisLeuValLeuLeuHisGlyValCysMetAlaGly---Asp 590

Db 2190 AGACAGGTTTCCCAACACATATAGTGTACTTACGGCGTGTGTGTCCGAGATGTGGA 2249

QY 591 SerIleMetValGlnGluPheValTyrLeuGlyAlaIaLeaspMetTyrLeuArgLysArg 610

Db 2250 AATATCATGGTGAAGAGTTGTGTGAGGGGGCGCGTGGATCTCTTCATGACCCGAGAA 2309

QY 611 GlyHisLeuValSerAlaSerTyrLysLeuGlnValThrLysGlnLeuAlaTyrAlaLeu 630

Db 2310 AGTATGCGGCTTACTACCCCTCGAAGTTCAAGGTGCGCAACACACTGCCAGTGGCCTG 2369

QY 631 AsnTyrLeuGlnAspLysGlyLeuProHisGlyAsnValSerAlaArgLysValLeuLeu 650

Db 2370 AGTTACTTGGAAAGTAAAGACTGGTTCTATGGAATATGTGCATTAAGAACCTCTCTTG 2429

QY 651 AlaArgGlnGlyLysAspGlyAsn---ProProPheIleLysLeuSerAspProGlyVal 669

Db 2430 GCCCGTAGGGCATGTGACAGTACATTTGGCCCGCTTCATCAAGCTTAGTGACCCGTGGATC 2489

QY 670 SerProThrValLeuSerLeuGluMetLeuThrAspArgIleProTyrValAlaProGlu 689

Db 2490 CCAATCTCTGTGCGTGAACAGGAGATGATAGAGCAATCCCTGGATGCGTCTGAG 2549

QY 690 CysLeuGlnGlnValaGlnThrLeuCysLeuGlnAlaAspLysTyrProGlyPheGlyAlaThr 709

D	2550	TCGTTTGAAGACTCCAGAAAGCCTGAGTGTGGCTGCTGACAAAGTGGAGCTTTGGAAACACAG	2603
Q	710	ThrTrpGluValPheGlnArgGlyProAlaHisLeuThrSerLeuGlnProAlaLysLys	729
D	2610	CTCTGGGAAATCTGCTCAACAGCGAGAGATCTCTCAAGAACAGACCCCTATTGTAGAA	2669
Q	730	LeuLysPheTyrGlnAspArgGlnGlyGlnLeuProAlaLeuLysTrpPheGlnLeuAlaLeu	749
D	2670	GAGAGGTTTATGAAAGCCCGCTGACGGCTGTGACTTCATCTTGCAAGAGAGCTTACTAC	2729
Q	750	LeuLeuThrGlnCysMetAlaTyrAspProGlyArgArgProSerPheArgAlaLeu	769
D	2730	CTCATGACTCGCTCATGACACTGAACTGAAACCCACAGACAGACCTTCTCTCGAGACCATATG	2789
Q	770	ArgAspLeuAsnGlyLeuLeuThrSerAspTyrGlnLeuLeuSerAspProThrProGly	789
D	2790	AGGGACATTAAACAAGCTCGAGAGAGACAGAAATCCAGCATTTGTTACAGAA	2837
Q	790	IleProSerProAlaArgAspGlnLeuCysValAlaGlyAlaGlnLeuTyrAlaCysGlnAsp	809
D	2838	---AAGCAGCCCAACACACAGAGTG-----GAC	2861
Q	810	ProAlaIlePheGlnGluArgHisLeuLysTyrIleSerLeuLeuGlyLysGlyAsnPhe	829
D	2862	CCCACTACACTTTGAAAGCGGCTTCGTAAGAGAGATTCTGACTTGCGAGAGAGGTACTTT	2921
Q	830	GlySerValGlnLeuCysArgTyrAspProLeuGlyAspAsnThrGlyProLeuValAla	849
D	2922	GGGAGAGGTGAGCTCTCAGATATGATCTCGAGGAGACACACAGGAGGAGCTACTCT	2981
Q	850	ValLysGlnLeuGln---HisSerValProAspGlnArgAspPheGlnArgGluLe	868
D	2982	GTCAAGTCCCTGGAAGCCTGAGAGTGGAGGTAAACACATAGCTGATCTGAAGAAGACATA	3041
Q	869	GlnIleLeuLysAlaLeuHisSerAspPheIleValLysTyrArgGlyValSerTyrGly	888
D	3042	GAGATCTTACGAGACCTCTACATGAGAACATTGTAAATACAAAGAAATCTCATGAA	3102
Q	889	ProGlyArgGlnSerLeuArgLeuValMetGluTyrLeuProSerGlyCysLeuArgAsp	908
D	3102	GACGAGAGCATGTGATCAAGCTCATACATGAGATGTTCTCTCGGAGAACTCTAAAGAG	3161
Q	909	LeuLeuGlnArgHisArgGly---LeuHisThrAspArgLeuLeuLeuPheAlaTrpGln	927
D	3162	TATCTGCCAAAGATTAAGAAACAAATCACTCAACACAGACTAAATATGCAATCCAG	3221
Q	928	IleCysLysGlyMetGluTyrLeuGlnAlaValArgArgCysValHisArgAspLeuAla	947
D	3222	ATTGTGAAGGAGAGACTACTTGGGTCTCGGCAATACGTTTACACGGGACTTAGACAGA	3281
Q	948	ArgAsnIleLeuValGlnSerGlnAlaHisValLysIleAlaAspPheGlyLeuAlaLys	967
D	3282	AGAAATGTCCTGTGTGAGAGTGAAGCATCAAGTCAAGATGCGAGACTTTGGTTTAACAA	3341
Q	968	LeuLeuProLeuGlnLysAspTyrTyrValValAlaValGlnProGlyGlnSerProIlePhe	987
D	3342	GCAATTGAACCCGATAGAGGTACTACAGAGTCAGAGACGACCGGAGACGCCACGTGTC	3401
Q	988	TrpTyrAlaProGlnSerLeuSerAspAsnIlePheSerArgGlnSerAspValTrpSer	1007
D	3402	TGGTAGCGTCCGGAGATGTTTAAATCCAGTGAATATTTATGCGCTCTGATGTCTGGCT	3461
Q	1008	PheGlyValValLeuTyrGlnLeuPheThrTyrCysAspLysSerCysSerProSerAla	1027
D	3462	TTTGATGACACACTGCGACGACGACGTCCTACTTCTGATCGATCAATTTAATGCCATGGC	3521
Q	1028	GluPheLeuArgMetMetGlyProGluArgGlnGlyProProLeuCysArgLeuLeuGln	1047
D	3522	TTGTGTTCTGAAATATGATAGGCCCACTCATGGCCAGATGACAGTACACACGCGTTGTAG	3581
Q	1048	LeuLeuAlaGluGlnArgArgLeuProProProProThrTyrGlnProThrGlnValGlnGlu	1067
D	3582	ACTGTGAAGAAAGGAAAGCGCTGTGGCATGTCACACCACTGTCTGATGAGAGGTTTATCAG	3641

```

Oy      1068 LemMeGlnLeuCytTrpAlaProGluPrnHisAspArgProAlaPhealThrLeu 1086
Db      3642 CTTATGACAAAGTGGTGGAATTCCCAACCATTACCGGACAACCTTTTCAGAACCTT 3698

RESULT 13
AAQ85413
ID      AAQ85413 standard; DNA; 3429 BP.
AC      XX
XX      AAQ85413;
DT      05-OCT-1995 (first entry)
DE      Human JAK1 kinase coding sequence.
KM      JAK family; protein tyrosine kinase; cytokine receptor;
KW      phosphorylation; signal transduction; activation; ss.
OS      Homo sapiens.
FH       Location/Qualifiers
FT       1..3429
CDS      /product= JAK1_kinase

MO9503701-A.
PD      09-FEB-1995.
PF      29-JUL-1994;   94MO-US08676.
PR      29-JUL-1993;   93US-0097997.
PA      (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.
PI      Thle JN, Quelle FW, Silvennoinen O, Witthuhn BA;
DR      WPI; 1995-081950/11.
PX      P-PSDB; AAR70831.
PT      Inhibiting a cellular response to a cytokine by inhibiting Jak
PT      kinase - to treat diseases caused by excessive response to
PT      cytokine, e.g. erythrocytosis and other cellular proliferative
PT      diseases
PS      Claim 29; fig 2; 167pp; English.
XX      Inhibiting the activity of a jak kinase (pref. Jak1, Jak2, Jak3 or
XX      Tyk2) in a eukaryotic cell is claimed as a method of inhibiting the
XX      biological response of that cell to a cytokine (not IL-3 or
XX      erythropoietin). The present sequence (human JAK1 kinase) was
XX      published by Wilks et al. in Mol.Cell.Biol. 11: 2057-2065 (1991).
XX      It encodes a 1142 amino acid protein which includes the epitopic
XX      sequence TLIERKRFVESRCRPVPS (amino acids 786-804). Antibodies which
XX      selectively bind this epitope are able to bind JAK1 without
XX      interfering with the activity of the kinase. Such antibodies are
XX      claimed and are useful for detecting and extracting JAK1.
XX      Sequence 3429 BP; 1011 A; 786 C; 856 G; 776 T; 0 other;

Alignment Scores:
Pred. No.:    1,16e-135          Length:    3429
Score:        1904.00           Matches:    443
Percent Similarity: 53.66%     Conservative: 188
Best Local Similarity: 37.67%   Mismatches:   381
Query Match: 32.49%           Indels:    164
DB:           Gaps:            24

US-09-397-967-16 (1-1099) x AAQ85413 (1-3429)
Oy      17 SetLenSerSetSerGlnAlaGlyAlaLeuHisValLeu---LeuProArgGlyPro 35
Db      46 AACCTTGAGAGGCCGTGCACGAGGTGTAATCTTTATCTCTGACAGAGGAGCCC 105

```


Db 3172 ACTCTGACGACGCTGCTGACTTACTGTGATTCTAGTCCATGGCTTTGTTCCCTG 3231
 Qy Argmetwetiylprogluargluclylproleucysargleuileuileuala 1050
 Db 3232 AAATGATAGGCCCAACCCATGCGCAGATGACAGTACACAAACCTTTGATACGTTAA 3291
 Qy 1051 GlnGlyArgArgLeuProProProProThrCysProThrGlnValGlnGlnLeuMetGln 1070
 Db 3292 GAAGGAAACGCGCTGCGCTGCGCACTAACTGTCAGATGAGGTTTACAGCTTATGAGA 3351
 Qy 1071 Leucylstrpalaiprogliuprohisaspargproalaiphealaiphrleu 1086
 Db 3352 AAATGCTGGGAATTTCACACATCCATGCGACAGCTTTTCAGAACCTT 3399
 RESULT 14
 AAC66245
 ID AAC66245 standard; DNA; 3429 BP.
 AC AAC66245:
 XX 19-FEB-2001 (first entry)
 DT 19-FEB-2001 (first entry)
 DE Human Jak1 kinase nucleotide sequence.
 XX
 KW Jak3; kinase; cytokine; cellular response; inhibition; Jak1; ds;
 KM cell proliferation; erythrocytosis.
 XX
 OS Homo sapiens.
 XX
 PN U6136595-A.
 PD 24-OCT-2000.
 XX
 PF 18-JUN-1996; 96US-0665574.
 XX
 PR 29-JUL-1994; 94US-0282012.
 PR 29-JUL-1993; 93US-0097997.
 PR 09-SEP-1993; 93US-0118968.
 XX
 PA (SUD-) ST JUDE CHILDREN'S RES HOSPITAL.
 PI Silvenoinen O, Maltuhon BA, Ihle J;
 DR WPI: 2000-666080/67.
 DR P-PSDB; AAB35725.
 XX
 PT New DNA encoding Jak3 kinase is useful as cytokine regulator for
 PT treating cell proliferation -
 XX
 PS Disclosure; Fig 2; 100pp; English.
 XX
 CC This invention relates to DNA encoding a murine Jak3 protein. The amino
 CC acid sequence of the Jak3 protein is given in AAB35725. The Jak family
 CC of kinases are involved in the cellular response to the binding of
 CC cytokines to their respective receptors. Jak3 kinase mediated activation
 CC of some cytokines through their phosphorylation in response to
 CC cytokine-receptor binding. Inhibiting the activity of Jak3 kinase (at the
 CC nucleic acid level with antisense sequences or ribozymes, or at the
 CC protein level with antibodies, kinase inhibitors etc.) is used to treat,
 CC or diagnose, diseases caused by excessive secretion of certain cytokines,
 CC e.g. excessive cell proliferation such as erythrocytosis. Alternatively,
 CC the Jak3 protein and polynucleotide can be used to treat conditions
 CC associated with defective Jak3 activity. The DNA sequence can be used to
 CC produce recombinant Jak3 and this used to raise antibodies useful as
 CC specific inhibitors or to detect or isolate Jak3 without interfering
 CC with its enzymatic activity. The present sequence represents DNA encoding
 CC a Jak1 protein used in the isolation and characterisation of the Jak3
 CC protein of the invention.
 XX
 SO Sequence 3429 BP; 1011 A; 786 C; 856 G; 776 T; 0 other;

Alignment Scores:

1.16e-135

Length:

3429

Score: 1904.00 Matches: 443
 Percent Similarity: 53.66% Conservative: 188
 Best Local Similarity: 37.67% Mismatches: 381
 Query Match: 32.49% Indels: 164
 DB: 21 Gaps: 24
 US-09-397-967-16 (1-1099) x AAC66245 (1-3429)
 Qy 17 SerleuserSerSerGlnAlaGlyAlaLeuHisValLeu-----LeuProProArgGlyPro 35
 Db 46 AACCTGGAGGCCCTGAGCCAGAGGGGTGAGAGTGAATCTTCTATCTGTGCGAGAGGAGGCC 105
 Qy 36 GlyProProGlnArgLeuSerPheSerPheGlyAspTyrLeuAlaGlnAspLeuValThr 55
 Db 106 -----CTCCGGCTGGGCGAGTGAGAGTACAGAGTACAGAGAACTGTGATC 150
 Qy 56 ArgAlaAlaLysAlaCysGlyIleLeuProValTyrHisSerLeuPheAlaLeuAlaThr 75
 Db 151 AGGGCTGCACAGGATGCGGTATCTCTCTTCTCACAACACTTTTGCCCTGTATGAC 210
 Qy 76 GlnAspPheSerCysTrpPheProPheSerHisIlePheCysIleGlnAspValAspThr 95
 Db 211 GAGAACACCAAGCTCTGTGTATGCTCCAAATGCGACCATCACCCTGTGATGACAGATGTC 270
 Qy 96 GlnValLeuValTyrArgLeuArgPheTyrPheProAspTyrPheGlyLeu----- 112
 Db 271 CTCCGGCTCCACTACCGGATGAGGTTCATTTTACACCATTTGGCATGGAACACAGACAAAT 330
 Qy 113 ---GluThrCysHisArgPheGlyLeuArgLysAspLeu----- 124
 Db 331 GAGCAGTCAGTGTGCGCATTTCTCCAAAGAGAGAAATGCGTACGAGAAAAAAG 390
 Qy 125 -----ThrSerAlaIleLeuAspLeuHisValLeuGlnHisLeuPheAlaGlnHis 141
 Db 391 ATTCAGATGCAACCCCTCTCTGTATGCTCCAGCTCAGTACGAGATCTTTTGTCTCAGGGA 450
 Qy 142 ArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLysGln----- 159
 Db 451 CAGTATGATTTGGTGAATGCGCTGCTCTATTCAGACACCAAGACCGACAGAGATGGA 510
 Qy 160 -----GlyGluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetValArg 175
 Db 511 CATGATATTGAGAACGAGCTGTCTAGGATGCTGTGCGCATCTCACTATGCCATG 570
 Qy 176 GlnGlnAlaGlnArgProGlyGlnLeuLeuLysThrValSerTyrLysAlaCysLeuPro 195
 Db 571 ATGAAGAGATGACAGTTGCGCAAGACTGCCCAAGACATCAGCTACACAGATATATTCGA 630
 Qy 196 ProSerLeuArgAspValIleGlnGlyGlnAspPheValThrArgArgArgIle----- 213
 Db 631 GAACATTTGAATTAAGTCCATCAGACAGAGAACTTTCACACAGATCGGATTAATAT 690
 Qy 214 -----ArgArgThrValValLeuAlaLeuLeuPro 223
 Db 691 GTTTTCAAGCATTTCTTAAGAGATTTTAACAACAAGACATTTGTGACAGCAGCGTG--- 747
 Qy 224 CysGlyArgLeuProGlyArgProTyrAlaLeuMetAlaLysTyrIleLeuAspLeuGlu 243
 Db 748 -----TCCACGATGACCTCGAAGGATGAATATCTTGCTACCTTGGA 789
 Qy 244 ArgLeuHisProAlaAlaIleThrGlnThrPheArgValGlyLeuProGlyAlaGlnGlu 263
 Db 790 ACTTTCACAAACATTTAGCGTGTGTAATATTT----- 822
 Qy 264 GluProGlyLeuLeuArgValAlaGlnLysAspGlnGlyIleProTyr---SerSerAsnAsp 282
 Db 823 GAGACTTCAGATGTACTGATTTATCATCAGAAATGAGATGAATGTGTTTCATTCGAATAC 882
 Qy 282 ----- 282
 Db 883 GGTGAACAGTTTCTCTACTAGAGATGATGTTGACTGGGAATCTTGGAATCCAGTGGAG 942
 Qy 282 ----- 282

DB 943 CATAAACCAATGTGTCTTCTGTAAGAAAAAATAACTGAAGCGGAAAAAATCTG 1002
QY 283 -----GluLeuPheGlnThrPhe 288
DB 1003 GAAATTAAGACAGAAGATGAGGAGAAAAACAGATCCGGAGAGTGAACAATTTT 1062
QY 289 CysAspPheProGluLeuValAspValSerLLeasnglnAlaProArgValGlyProAla 308
DB 1063 TCATTCTCTCCCTGAATCACTCATCTGTAATTAAGAGACT----- 1104
QY 309 GlyLLeuHisArgLeuValThrValThrArgMetAspGlyHisLLeuGlnAlaGluPhe 328
DB 1105 -----GTGGCACCATTAAACACAGCAGACACACAGAAAGAAAGCACTCAACCTC 1152
QY 329 ProGlyLeuProGlnAlaLeuSerPheValAlaLeuValAspGlyTyrPheArgLeuIle 348
DB 1153 TCTTCCACAGAGAGCGCTGTCTTGTCTCCCTGTAGATGGCTACTCCGGCTCACA 1212
QY 349 CysAspSerArgHisTyrPheCysLysGlnValAlaProProArgLeuGlnGluGlu 368
DB 1213 GCAGATGCGCATATTAATCTCTGACCGAGCTGCCCCGTTGATGCTCACACACATA 1272
QY 369 AlaAspValCysHisGlyProIleThrLeuAspPheAlaIleHisLysLeuValAla 388
DB 1273 CAGAAATGGCTGCATGGTCCCATCTGTACAGAAATAGCCCATCAATTAATTTGGCGCAAGAA 1332
QY 389 GlySerLeuProGlyThrTyrIleLeuArgArgSerProGlnAspTyrAspSerPheLeu 408
DB 1333 GGAAGGAGAGGAGGATGATGCTGAGGTGAGCTGACACCACTTGACACACTCTC 1392
QY 409 LeuThr-----AlaCysValGlnThrPro-----LeuGlyProAspTyrLys 422
DB 1393 ATGACCGTCACTCCCTTGTGAGAGTGCAGACAGGTGCAGGGTGGCCGAGACAGCTTCAAG 1452
QY 423 GlyCysLeuIleArgGlnAspProSerGlyAlaPheSerLeuValGlyLeuSerGlnPro 442
DB 1453 AACCTTCAGATC---GAGGTGCAGAAAGGCGCGTACAGTCTGCACAGCTTGCAGCCGAGC 1509
QY 443 HisArgSerLeuArgGluLeuAlaAlaCysTrpAsnSerGlyLeuArgValAspGly 462
DB 1510 TTCCCCAGCGTTGGAGAGCTCTGATGAGCCACCTCAGACAGACAGATCCGCGACGAGTAAC 1569
QY 463 AlaAlaLeuTyrLeuThrSerCysAlaProArgProLysGlnLysSerAsnLeuIle 482
DB 1570 ATGAGCTTCATGCTAAAGCTGCTGCCACCAAGCCCGAGAAATCTTCACACCTCTC 1629
QY 483 ValValArgArgGlyCysAsnProAlaProAlaProGlyCysSerProSerCysAla 502
DB 1630 GTGGCTACTTAAG-----AAAGCCAGAGAGTGGCAGCCC---GTCTACCCC 1671
QY 503 LeuThrGlnLeuSerPheHisThrIleProThrAspSerLeuGluTyrPHisGlnAsnLeu 522
DB 1672 ATGAGCCAGCTGAGTTTCATCGGATCCTCAGAAAGATGTGGTGCAGGCGAGCACCTT 1731
QY 523 GlyHisGlySerPheThrLysIlePheArgGlySerArgArgLysValAlaAspGlyGlu 542
DB 1732 GGGAGAGGACGAGAAACACATCTATTTGGGACCTGTAGTGTACAGATGAGAGAA 1791
QY 543 ThrHisAspSerGlu-----ValLeuLeuLysValMetAspSerArgHisArg 558
DB 1792 GGAACCTTGAAGAGAAAGATTAAGATGATCTCAAGCTTTAGACCCACGACGAG 1851
QY 559 AsnCysMetGluSerPheLeuGlnAlaAlaSerLeuMetSerGlnLysTyrProHis 578
DB 1852 GATATTTCCCTGGCTTCTTCGAGGAGCGACAGATGTGAGACAGGCTCTCCACAAACAC 1911
QY 579 LeuValLeuLeuHisGlyValCysMetAlaGly---AspSerIleMetValGlnGluPhe 597
DB 1912 ATCGTGTACTATGCGCTGTGTGTCGAGAGCTGAGAAATATCATGTGTGAAGAGCTTT 1971
QY 598 ValTyrLeuGlnAlaIleAspMetTyrLeuArgLysArgGlyHisLeuValSerAlaSer 617
DB 1971 ----- 617

DB 1972 GTGGAAGGGGCTCCTCGATCTCTTCATGACCGGAAAAAGTATGTCCTTACACACCA 2031
QY 618 TrpLysLeuGlnValThrLysGlnLeuAlaTyrAlaLeuAsnTyrLeuGluAspLysGly 637
DB 2032 TGGAAATTCAAAGTTGGCAACAGCTGGCCAGTCCCTGTGACTTGGAGGATTAAGAC 2091
QY 638 LeuProHisGlyAsnValSerAlaArgLysValLeuLeuAlaArgGlnGlyLysGly 657
DB 2092 CTGGCTCATGAAATGTGTACTAAAACCTCTCTGCGCCGCTGAGGAAATCGACAGT 2151
QY 658 Asn---ProProPheIleLysLeuSerAspProGlyValSerProThrValLeuSerLeu 676
DB 2152 GAGTGTGGCCATTCATCAAGCTCAGTACGCCGCGCATCCCATACGAGTGTCTAGG 2211
QY 677 GluMetLeuThrAspArgIleProTyrPValAlaProGlyCysLeuGlnAlaGlnThr 696
DB 2212 CAAGAAATGATTAAGAACCAATCCATGATGTCCTGAGTGTGAGGACTCAAAAC 2271
QY 697 LeuCysLeuGlnAlaAspLysTyrPValPheGlyValThrThrTrpGluValPheGlnArg 716
DB 2272 CTGAGTGTGGCTGTGACAGTGTGAGCTTGGAAACCGCTCTGGGAAATCTGCTACAT 2331
QY 717 GlyProAlaHisIleThrSerLeuGlnProAlaLysLysLeuLysPheTyrGluAspGln 736
DB 2332 GCGGAGATCCCTTGAAGAGCAAGACCGCTGATGAGAAAGAGATCTATGAAGCCG 2391
QY 737 GlyGlnLeuProAlaLeuLysTrpThrGlnLeuAlaGlyLeuLeuThrGlnCysMetAla 756
DB 2392 TGCAGCGCATGACACCATATGATAGAGACTGTGCTACTGACCTGACCTGCTGATGAC 2451
QY 757 TyrAspProGlyArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIle 776
DB 2452 TATGACCCCAATCAGAGAGCTTCTTCCGAGCATCATGAGACATATTAATGCTTGA 2511
QY 777 ThrSerAspTyrGluLeuLeuSer-----AspProThrProGlyIleProSer 792
DB 2512 GAGCAGATCCATGATATGTTCCAGAAAAAACCAGCCACTGAA----- 2559
QY 793 ProArgAspGluLeuCysValAlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIle 812
DB 2560 -----GTGACCCCAACCAT 2574
QY 813 PheGlnGlnArgHisLeuLysTyrIleSerLeuLeuGlyLysGlnAsnPheGlySerVal 832
DB 2575 TTGAGAGAGCGCTTCCCTAAAGAGAGATCCGATGCTGGAGAGGCGCTTGGGAAGGTT 2634
QY 833 GluLeuCysArgTyrAspProLeuGlyAspAsnThrGlyProLeuValAlaValLysGln 852
DB 2635 GAGCTGTGAGGTATGACCCC---GAAGACAAATACAGGAGGAGGCTGTAAATCT 2691
QY 853 LeuGln---HisSerValProAspGlnArgAspPheGlnArgLysIleGlnIleLeu 871
DB 2692 CTGAGCGCTGAGAGTGGAGGTAAACCATAGCTGATGTAAGAAAGAAATCGAGATCTTA 2751
QY 872 LysAlaLeuHisSerAspPheIleValLysTyrArgGlyValSerTyrGlyProGlyArg 891
DB 2752 AGGAACTCTATCATGAGCAATGTGAATACAAAGAAATCTGCACAGAGAGCGGGA 2811
QY 892 GlnSerLeuArgLeuValMetGluTyrLeuProSerGlyLysLeuArgAspLeuGln 911
DB 2812 AATGCTATTAGCTCATCATGGAATTTCTGCTTGGGAAGCCTTAAGAAATCTTCCA 2871
QY 912 ArgHisArgGly---LeuHisThrAspArgLeuLeuPheAlaTrpGlnIleCysLys 930
DB 2872 AAGATTAAGAAACAAATAAATCAACACAGCTGAATAATATGCGCTTCAGATTTGTAAG 2931
QY 931 GlyMetGluTyrLeuGlnGlyAlaArgArgCysValHisArgAspLeuAlaIleArgAsnIle 950
DB 2932 GGGATGACATATTTGGTCTTCGCAATACCTTCACCGGAGACTTGGCAGAGAAAGTGC 2991
QY 951 LeuValGlnSerGluAlaHisValLysIleAlaAspPheGlyLeuAlaLysLeuLeuPro 970
DB 2992 CTGTGTGAGGTGAAACCAAGTGAATGAGACTTGGCTTTAACCAAGCAATTTGAA 3051

Oy	971	LeuGlyLysAspArgTyrValValArgGluProGlyGlnSerProIlePheTrpTyrAla	990
Db	3052	ACCCATTAAGGAGTATTACACCGTCAGAGATGACCGGACACCCCTGTGTGGTATGCT	3111
Oy	991	ProGluSerLeuSerAspAsnIlePheSerArgGlnSerAspValTyrSerPheGlyVal	1010
Db	3112	CCAGATAGTTTAATGCAATCTTAATTTTATATTGTCCTGACGCTGTGCTTTGGAGATC	3171
Oy	1011	ValLeuTyrArgLeuLeuPheThrTyrCysAspLysSerCysSerProSerAlaGluPheLeu	1030
Db	3172	ACTCTGCATGAGCTGCCTGACTTACTGTGATTCAGATTCTAGTCCCATGGCTTGTCTCG	3231
Oy	1031	ArgMetMetGlyProGluArgGlyGlyProProLeuGlySerArgLeuLeuGluLeuAla	1050
Db	3232	AAATGATAGGCCCAACCCATGCGCCAGATGACAGTCAACAGACTTGCATACGTTTAAA	3291
Oy	1051	GluGlyArgArgTyrGluProProProProTyrCysProThrGluValGlnGluLeuMetGln	1070
Db	3292	GAAAGGAAACGCCCTGCGCTGCCACCTTAACCTGCACATGAGCGTTTATGACGTTATGACA	3351
Oy	1071	LeuGlySerPheAlaProGluProIleHisAspArgProAlaPheAlaThrLeu	1086
Db	3352	AAATGCTGGGAATTCCAACCATCCAAATCGGACCAAGCTTTCAGAACTT	3399
RESULT 15			
ID	AAD03608	standard; cDNA; 3429 BP.	
AC	AAD03608;		
XX			
AC	AAD03608;		
XX			
DT	19-JUN-2001	(first entry)	
XX			
DE	Human (Janus kinase 1) Jak1 kinase cDNA.		
KW	Human; Janus kinase 1; Jak1; therapy; cytokine; cytokine; cellular response;		
KW	antiproliferative; cytosstatic; cell proliferative disorder;		
XX	tyrosine kinase; ss.		
OS	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	1..3429	
FT		/*tag= a	
FT		/product= "Human (Janus kinase 1) Jak1 kinase"	
XX			
PN	US6210654-B1.		
XX			
PD	03-APR-2001.		
XX			
PF	08-OCT-1997;	97US-0946994.	
XX			
PR	18-JUN-1996;	96US-0665574.	
XX			
PR	29-JUL-1993;	93US-0097997.	
XX			
PA	(SJUD-) ST JUDE CHILDREN'S HOSPITAL.		
XX			
PI	Ihle J, Witthuhn BA, Quelle FW, Silvennoinen O;		
XX			
DR	WPI: 2001-265367/27.		
XX			
DR	P-PSDB: AAE00353.		
XX			
PT	Modulating a biological response mediated by Jak kinase 2 activation to		
PT	a cytokine, useful for treating excessive proliferation of eukaryotic		
PT	cells, comprises inhibiting or enhancing tyrosine kinase activity of		
PT	Jak kinase in the cell -		
XX			
PS	Example 1: Fig 2: 100pp: English.		
XX			
XX	The present sequence is a cDNA encoding human (Janus kinase 1) Jak1		
CC	tyrosine kinase. Jak kinases mediate cytokine activity through their		
CC	tyrosine phosphorylation in response to cytokine-receptor binding.		
CC	The present invention relates to a method for modulating response of		

CC eukaryotic cells to a cytokine by regulating the tyrosine kinase activity
CC of Jak kinases. This method is useful in treating diseases caused by an
CC excessive cellular response to a Jak kinase mediated cytokine, by
CC inhibiting Jak kinase activity. Such diseases include those caused by
CC excessive proliferation of cells. Treatment of disease conditions caused
CC by a deficient cellular response or non-responsiveness to a cytokine
CC involves enhancing Jak kinase activity. Jak kinase sequence is useful
CC in identifying its inhibitors.

Sequence 3429 BP; 1011 A; 786 C; 856 G; 776 T; 0 other

Alignment Scores:

Pred. No.:	1.16e-135	Length:	3429
Score:	1904.00	Matches:	443

Percent similarity:	53.668	conservative:	188
Best local similarity:	37.678	Mismatches:	381

Query Match:	32.498	Indels:	164
DB:	22	Cases:	24

115-00-307-067-16 (1,000) 2,220,000 1,000

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

```
x7      ::|||   |||   ::|||
         ::|||   |||   ::|||
         ::|||   |||   ::|||
```

40 AACCTGGAGGCCCTGAGCCAGGGCTGGAAGTGTCTTCTATCTGTG

36 GlyProGlnArgLeuSerPheSerPheGlyAspTyrLeuAlaGlu

Db 106 -----CTCCGGCTGGGCAGTGGAGGTACACAGCAGAG

Qy 56 ArgAlaAlaLysAlaCysGlyIleLeuProValTyrHisSerLeuPhe

[illegible]

Ov 76 G]wAsrBhOSeRClucmnhDnDndGcui - rj - n! - c - u] - e]

[illegible][illegible][illegible]

2/1 C TCCGGC TCCACTACCGGATGAGGTTC TATTTCCACCAATTGGCATGGA

QY 113 ---GluThrCysHisArgPheGluLeuArgLysAspLeu-----

Db 331 GAGCAGTCAGTGTGGCTCATTTCTCCAAGAGCAGAAATGGCTACC

QY 125 -----ThrSerAlaIleLeuAspLeuHisValLeuGluHisLeu

Db
391 ATTCAGATGCAACCCCTCTTCTTGATGCGAGCCTACTGCACCTATCTCG

[illegible][illegible][illegible]

100 -----GlyciupneuserLeuAlaValLeuaspleuAlac
||| :||| :
||| :||| :

511 CATTGATTTGACACGAGTGTCTAGGATGGCTGTCCCTGGCCATCTCA

QY 176 G I U G I N A I A G I n A r g P r o g I y G I U L e u L e u L y s t h r V a l S e r T y r L y s

Db 571 ATGAGAGATGCAGTTGCCAGACTGCCCAAGACATCAGCTACAAGC

196 ProSerLeuArgAspValIleGlnIvcInasnpheValThrArqArqA

631 GAACATTTGATTAAGTTCATCAGACAGAGCACTTCTCTACCACTTCT

214

[illegible][illegible]

227 cy501yagpearfio01yagfio1yralaleumetalalystyrllel
::: ||| |||:::

DD /48 -----TTCACGCATGACCTGAAGGTGAATACTTGG

CCITGGGA 789

Oy	244	ArgLeuHisProAlaIaIaThrThrgluThrPheArGValIGlyLeuProGlyAlaGlnGlu	263
Db	790	ACTTTGCAGAAACATTACGGTGTGGTAATATTT-----	822
Oy	264	GluProGlyLeuLeuArGValAlaGlyAspAnGlyLleProTrp---SerSerAsnAsp	282
Db	823	GAGACTTCACGTtAcTGATGATTCATCCAGAAAATGAGATGAATGGTTTCATTTCGAATGAC	882
Oy	282	-----	282
Db	883	GGTGGAAACCTTCTCACTAACAGAAAGTATGGTACGGGAATCTTGAATCAGTGGAGG	942
Oy	282	-----	282
Db	943	CATAAACCATAATGTTGTTCTGTGAAAAAGAAAAAATAATAACTGAAAGCGAAAAAATCG	1002
Oy	283	-----GluLeuPheGlnThrPhe	288
Db	1003	GAAATTAAGACACAAGAAGATGAGAGAAAAAACAAGATCCGGAGAGATGGAGACAATTTT	1062
Oy	289	CysAspPheProGluLlEvaLasrValSerLlleasnGlnAlaProArGValGlyProAla	308
Db	1063	TCAATTTCCCTCGAAACACTCACATTTGATTAAGAAAGACT- - - - -	1104
Oy	309	GlyGluHisArGLeuValAlaThrValAlaThrArMetLasrGlyHisLlleuGlnAlaGluPhe	328
Db	1105	- - - - -GTGTCACACATTAAACAAGACAGACACAAACAAACAAATAAGTGAACCTGAAGCTC	1152
Oy	329	ProGlyLeuProGluAlaIleuSerPheValAlaIleuValAsrGlyTYrPheArGLeuLle	348
Db	1153	TTTTCCCAACGAGAGGCCCTTGCTTTGTGTGCTCTGTAAGTGGTACTTCGGCTGCACA	1212
Oy	349	CysAspSerArGHisTyrrPheCysLysGluValAlaIaIaProArGLeuLeuGluGlu	368
Db	1213	CGAGATGCCCATTCATTAACCTGTGCACCGACGAGCTGGCCCCCTGTGATCGTCCACAACTA	1272
Oy	369	AlaAsrValCysHisGlyProLlleThrLeuAsrPheAlaLlleHisTyLeuLysAlaIa	388
Db	1273	CAGAATGGCTCTCATATGGTCACATTTGACAGAAATAGCCCATTAATAATTGGCGCAAGA	1332
Oy	389	GlySerLeuProGlyThrTyrlleLeuArGArGserProGlnAspTYrAspSerPheLeu	408
Db	1333	GGAACGAGAGGGGATGTACGTCTGAGGGGAGCTGACACGACTTTGACAACTCTCTC	1392
Oy	409	Leuthr---AlaCysValGlnThrPro-----LeuGlyProAspTYrLys	422
Db	1393	ATGACCGTCACCTGCTTTGAGAAAGTGTGACAGAGTGCAGGGTCCCAGAAAGCAATTCAAG	1452
Oy	423	GlyCysLeuLleArGlnAsrProSerGlyAlaPheSerLeuValGlyLeuSerGlnPro	442
Db	1453	AACTTTCAGAT--GAGGTGCAGAAAGGCGCGTAGACGTTCAGCGTTCCGACGCGCAGC	1508
Oy	443	HisArGserLeuArGluLeuLeuAlaIaIaCysTrpAsnSerLlleuArGValAsrGly	462
Db	1510	TTTCCCAAGCTTGCGGACCTCATGAGCCACCTCAAGAACAGATCTCTGGACAGGATAAC	1566
Oy	463	AlaIaLeuTyrrLeuThSerCysCysAlaProArGProLySGlyLysSerAsnLeuLle	482
Db	1570	ATCAGCTTATCTAANAACGCGTGGCCAGCCCAAGCCCGAGAAATCTCAACCTGTG	1626
Oy	483	ValAlaArGArGlyCysAsnProAlaProArLleProGlyCysSerProSerCysCysAla	502
Db	1630	GTGGCTACTAAG-----AAACCCAGAGAGTGGCAGGCC--GTCTACCCC	1677
Oy	503	LeuThrGlnLeuSerPheHisThrLleProThrAsrPheLeuGlnTyrrHisGluAsnLeu	522
Db	1672	ATGAGCCACTAGTTTGCATGGATCGATCTCCAAGAAAGATCTGGTGCAGGGCGAGACCTT	1733
Oy	523	GlnHisGlySerPheThrTyrrLlePheArGlyLleSerArGArGValAlaAsrGlyGlu	542
Db	1732	GGGAGAGGAGCAGAAACACATCTATCTGGAGCCCTGATGAGCTTCAAGAGATGACGAA	1799
Oy	543	ThrHisAsrSerGlu-----ValLeuLeuLysValImeTarSerArGHisArG	558

Db	1792	GGAAGCTTCTAAGAGAGAGAGATAAAGATGATCCCTCAAGAGCTTTAAGACCCAGCCAGG	1851
Qy	559	AsnGlyMetGlnSerPheLeuGlnAlaAlaSerLeuMetSerGlnValSerTyrProHis	578
Db	1852	GATATTTTCCCTGGCCCTTTCCTCGAGGAGCCACACATGATGAGACAGCTTCCCAAAAC	1911
Qy	579	LeuValLeuLeuHisGlnValAlaGlyMetAlaGly---AspSerIleMetValGlnIlePhe	597
Db	1912	ATGCGTAGACTCTATGAGCGCTGTGTGTCGCCGACGATGAGAAATATCATGTGTGGAAGATT	1971
Qy	598	ValTyrLeuGlnAlaAlaLeuAspMetTyrLeuAlaGlyValGlyHisLeuValSerAlaSer	617
Db	1972	GTGGAAGAGGGATCCCTCGATCTCTTCATGCGACACGGAAAGATGATCTTTACACACCA	2031
Qy	618	TrypLeuLeuGlnValAlaThrTyrGlnLeuAlaTyrAlaLeuAsnTyrLeuGlnAspGlyGly	637
Db	2032	TGGAAATTCAAAGTGGCCAAACAGCTGGCCAGTGGCCCTGAGCTACTTGTGAGGATAAAGAC	2091
Qy	638	LeuProHisGlnValSerAlaArgLysValLeuLeuAlaArgGlnGlyLysProGly	657
Db	2092	CTGGCTCCATGGAATGTGTACTATAAAACCTTCCTCTGGCCGAGAGGATCGACAGT	2151
Qy	658	Asn---ProProPheIleLysLeuSerAspProGlyValSerProThrValLeuSerLeu	676
Db	2152	GAGTGTGGCCATTATCATCAGCTCAGTGAACCCGGCATCCCATTAAGGTGTGTCTAG	2211
Qy	677	GlnMetLeuThrAspArgIleProTyrValAlaProGlnLysLeuGlnGlnValAlaThr	696
Db	2212	CAGAAATGCATTGAAACAAATCCCATGATGTGCTCCGAGAGTGTGTGAGGATCCCAAGAC	2271
Qy	697	LeuGlyLeuGlnValAspLysTyrGlyPheGlnValAlaThrTyrProLysValPheGlnArg	716
Db	2272	CTGAGTGTGTGGCTGTGAAGATGAGAGCTTTGGAACACCGCTGGAATCTGCTACAAAT	2331
Qy	717	GlyProAlaHisIleThrSerLeuGlnProAlaLysLeuLysLeuLysPheTyrGlnLysGln	736
Db	2332	GGCGAGATCCCTTGAAAGACAAACAGCCTGATTGAGAAAGAGATTCATATAAGCCGG	2391
Qy	737	GlnLysLeuProAlaLeuLysTyrThrGlnLeuAlaGlyLeuIleThrGlnLysMetAla	756
Db	2392	TGCAGGACAGTGAACACCTCATGTAAGAGCTGGCTGACCTCATGACCCGTCGATGAAC	2451
Qy	757	TyrAspProGlyArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIle	776
Db	2452	TATGACCCCAATCAAGAGCCCTTCTTCCGAGCATCATGAGACATCTAATAAGCTGAA	2511
Qy	777	ThrSerAspTyrGlnLeuLeuSer-----AspProThrProGlyIleProSer	792
Db	2512	GAGCAGAAATCCAGATATTGTTTCCAGAAAAAAAACAGCCAACTGAA-----	2555
Qy	793	ProArgAspGlnLeuGlyValAlaGlnLeuTyrAlaLysGlnAspProAlaIle	812
Db	2560	-----GTGGACCCCAACAT-----	2577
Qy	813	PheGlnGlnArgHisLeuLysTyrIleSerLeuLeuGlnLysGlnLysAsnPheGlySerVal	832
Db	2575	TTTGTGAACCGCTTCTTAAGAGATCCGTACTTGGAGAGGCCACTTGGGAAGGTT	2633
Qy	833	GlnLeuGlyArgTyrAspProLeuGlnLysAsnThrGlyProLeuValAlaValLysGln	852
Db	2635	GAGCTGTGACGATAGACCC---GAAACATATCAGAGGAGCGAGGCGCTGTTAAATCT	2699
Qy	853	LeuGln---HisSerValProAspGlnIleArgAspPheGlnArgGlnIleGlnIleLeu	871
Db	2692	CTGAACCTCGAAGTGGAGGTACCAACATACCTATCTGAAAAAGGAATCGAGATCTTA	2755
Qy	872	LysAlaIleHisSerAspPheIleValLysTyrIleArgLysIleSerTyrGlnProGlyArg	891
Db	2752	AGGAACCTCTATCATGAGACACATGTGTGAAGTAAAGAAATCTGCACAGAAAGCGAGGA	2811
Qy	892	GlnSerLeuArgLeuValMetGlnTyrLeuProSerGlyLysLeuArgAspLeuLeuGln	911

Db 2812 AATGATATTAGCATCATGATGATTTCTCTCGGAGACCTTAAGGATATCTTCA 2871
QY 912 ArgHisArgGly--LeuHisThrAspArgLeuLeuLeuPheAlaTrpGlnIleCysLys 930
Db 2872 AAGATATAGAACAAATTAACCTCAACACACACTTAATATAGCCGTTGAGATTGTAG 2931
QY 931 GlyMetGluTyrLeuGlyAlaArgArgCysValHisArgAspLeuAlaAlaArgAsnIle 950
Db 2932 GGGATGGACTATTGGGTTCTCGCAATACCTTACCCGGGACTTGGCAGCACAAGATGTC 2991
QY 951 LeuValGluSerGluAlaHisValLysIleAlaAspPheGlyLeuAlaLysLeuLeuPro 970
Db 2992 CTGTGTCAGAGTGACCAACCAAGTGAATTTGAGACTTCGTTTAACCAAGCAATTGAA 3051
QY 971 LeuGlyLysAspTyrTyrValValArgGluProGlyGlnSerProIlePheTrpTyrAla 990
Db 3052 ACCGATTAAGAGATTATACACCGTCAGAGATGACCGGACAGCCCTGTGTGTTGGTATGCT 3111
QY 991 ProGluSerLeuSerAspAsnIlePheSerArgGlnSerAspValTrpSerPheGlyVal 1010
Db 3112 CCAGATGTTTATGCAATCTAATTTTATATGCTGACCTGACCTGTGCTTTGGAGTC 3171
QY 1011 ValLeuTyrGluLeuPheThrTyrCysAspLysSerCysSerProSerAlaGluPheLeu 1030
Db 3172 ACTTGCATGAGCTGCTGACTTACTTACTGATTCTAGATTCTAGTCCATGGCTTGTTCCTG 3231
QY 1031 ArgMetMetGlyProGluArgGluGlyProProLeuCysArgLeuLeuGluLeuAla 1050
Db 3232 AAATGATAGGCCCAACCCATGCCAGATGCACAGTCAACAAGACTTGTGAATACGTTAAAA 3291
QY 1051 GluGlyArgArgLeuProProProProThrCysProThrGluValGlnGluLeuMetGln 1070
Db 3292 GAAGGAAACGCGCTGCCGTCACCTAAGTGCAGATGAGGCTTATCAGCTTATGAGAG 3351
QY 1071 LeuCysTrpAlaProGluProHisAspArgProAlaPheAlaThrLeu 1086
Db 3352 AATGCTGGGAATTCACACCATCCAAATCGGACAGCTTTCAGAACCTT 3399

Search completed: April 28, 2003, 14:03:09
Job time : 735.936 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.5
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: April 28, 2003, 16:07:34 ; Search time 302.323 Seconds
(without alignments)
3955.543 Million cell updates/sec

Title: US-09-397-967-16

Perfect score: 5860
Sequence: 1 MAPSEETPLIPORCSLSS.....RPAFATLSPQLDPLWRGRPG 1099

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 709820 seqs, 544064369 residues

Total number of hits satisfying chosen parameters: 1419640

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Command line parameters:
-MODE=frame_plus.model -DEV=xlp
-Q=/cgn2_1/USPTO_SPOOL/US09397967/Runat_26042003_182314_10830/app.query.fasta_1.1486
-DB=Published_Applications_NA -QMT=fastap -SUFFIX=rpnb -MINMATCH=0.1
-LOOPT=0 -LOOPTXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=BLOSUM62
-TRANS=human40.cdi -LIST=45 -DOCLIGN=200 -THR_SCORE=pcct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEPSSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09397967@cgn_1_1_202_@runat_26042003_182314_10830
-NCPU=6 -ICPU=3 -NO_XLERY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT
-DSPBLOCK=100 -LONSTOG -DEV_TIMECUT=120 -WARN_TIMECUT=30 -THRADS=1 -XGAPOP=10
-XGAPEXT=0.5 -FCGAPOP=6 -FCGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: Published_Applications_NA

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCIT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCITUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1908.5	32.6	4191	9	US-09-962-854A-1
2	1904	32.5	3541	10	US-09-880-107-2379
3	1842	31.4	3538	9	US-09-962-854A-2
4	713	12.2	2338	10	US-09-925-302-90

Sequence	Score	Length	Matches	Conservative	Mismatches
Sequence 29, Appl	609.5	10.4	2027	10	US-09-771-161A-29
Sequence 5, Appl	483.5	8.3	975	9	US-09-962-854A-5
Sequence 636, App	482.5	8.2	753	10	US-09-910-943-636
Sequence 148, App	482	8.2	3914	12	US-10-044-090-148
Sequence 8, Appl	475	8.1	3768	10	US-09-811-123-8
Sequence 2, Appl	475	8.1	3768	10	US-09-811-115-2
Sequence 11, Appl	475	8.1	4530	9	US-09-877-177-11
Sequence 7, Appl	475	8.1	9274	10	US-09-811-123-7
Sequence 1, Appl	475	8.1	9274	10	US-09-811-115-1
Sequence 9, Appl	474	8.1	3768	9	US-09-854-356-9
Sequence 1, Appl	474	8.1	3768	9	US-09-930-125-1
Sequence 5, Appl	474	8.1	4473	9	US-09-441-411-5
Sequence 1, Appl	474	8.1	4543	9	US-09-765-508-1
Sequence 1, Appl	465.5	8.0	3633	10	US-09-725-433-1
Sequence 98, Appl	465.5	7.9	5484	10	US-09-940-101-1
Sequence 10, Appl	464.5	7.9	10058	9	US-09-974-298-98
Sequence 117, App	464.5	7.9	3955	9	US-09-854-356-10
Sequence 19, Appl	464.5	7.9	3955	9	US-09-870-759-117
Sequence 48, Appl	455	7.8	3771	9	US-09-854-356-11
Sequence 271, App	455	7.8	7607	10	US-09-962-610-19
Sequence 3, Appl	450.5	7.7	3726	10	US-09-925-302-271
Sequence 48, Appl	448.5	7.7	2437	9	US-09-158-722-3
Sequence 1631, Ap	448.5	7.7	3454	12	US-10-044-090-48
Sequence 6, Appl	442.5	7.6	2863	10	US-09-954-456-1631
Sequence 4, Appl	441	7.5	1755	9	US-09-930-125-6
Sequence 7, Appl	441	7.5	1767	9	US-09-930-125-4
Sequence 5, Appl	441	7.5	1773	9	US-09-930-125-7
Sequence 1731, Ap	438	7.5	5264	10	US-09-930-125-5
Sequence 5, Appl	438	7.5	5264	12	US-10-033-528-1731
Sequence 1731, Ap	437.5	7.5	2770	9	US-09-927-260-5
Sequence 5, Appl	437.5	7.5	2770	9	US-09-977-261-5
Sequence 13, Appl	437.5	7.5	2770	10	US-09-977-269-5
Sequence 120, App	428.5	7.3	4879	9	US-10-172-620-13
Sequence 313, App	428.5	7.3	4975	10	US-09-880-107-2342
Sequence 3, Appl	425	7.3	2454	10	US-09-964-824A-120
Sequence 6, Appl	425	7.3	2627	10	US-09-964-436-313
Sequence 3, Appl	420.5	7.2	4541	9	US-10-123-036-3
Sequence 3785, Ap	419.5	7.2	3060	10	US-09-880-107-3785
Sequence 1, Appl	416	7.1	2674	9	US-09-824-859A-6
Sequence 1, Appl	416	7.1	2674	10	US-10-003-295-1

ALIGNMENTS

RESULT 1
US-09-962-854A-1
Sequence 1, Application US/09962854A
Publication No. US20030044803A1
GENERAL INFORMATION:
APPLICANT: Pedersen, Finn S.
APPLICANT: Hernandez, Javier Martin
TITLE OR INVENTION: Methods for diagnosis and treatment of diseases associated with
FILE REFERENCE: A-70020/RMS/DCF
CURRENT APPLICATION NUMBER: US/09/962,854A
PRIOR FILING DATE: 2001-09-24
PRIOR APPLICATION NUMBER: US 09/668,644
NUMBER OF SEQ ID NOS: 5
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 4191
TYPE: DNA
ORGANISM: Mus sp.
US-09-962-854A-1

Alignment Scores:
Pred. No.: 2,29e-191
Score: 1908.50
Percent Similarity: 54.52%
Best Local Similarity: 38.37%

Length: 4191
Matches: 437
Conservative: 184
Mismatches: 383

Db	2430	GCCCGTGGGGCATTGACACTGACATTGGCCGCTTCATCACTTAAGTACCCCTGGCATC	2489
Oy	670	SerProThrValLeuSerLeuGluMetLeuThrAspArgIleProTrpValAlaProGlu	689
Db	2490	CCAGTCTCTGGTGGACCCAGCAGAAAGATGCATAGACCGCATCCCTGGATGCTCTCTAG	2549
Oy	690	CysLeuGlnGlnAlaGlnThrLeuCysLeuGlnAlaAspLysTrpGlyPheGlyValAlaThr	709
Db	2550	TGCTTGACAGCTCCAGAAACCTGAGTGTGGCTGCTGCAGACAGTGGAGCTTTGGAAACACG	2609
Oy	710	ThrTrpGluValPheGlnArgGlyProAlaHisIleThrSerLeuGluProAlaLysLys	729
Db	2610	CTCTGGAGAAATCTGCTCAACAGGAGATCCCTCTCAAGACAAAGACCCCTATTGGANA	2669
Oy	730	LeuLysPheArgValAspGlnGlnGlnLeuProAlaLeuLysTrpThrGlnLeuAlaGly	749
Db	2670	GAGAGGTTTATGAAAGCCGCTCCAGGCGCTGTACTCCATCTTGCAAGAGAGCTAGCTGAC	2729
Oy	750	LeuIleThrGlnCysMetAlaTrpAspProGlyArgArgProSerPheArgAlaIleLeu	769
Db	2730	CTCATGACTGCGTGCATGAACTATGACCCCAACAGAGACCTTCTCCAGGACATATG	2789
Oy	770	ArgAspLeuAsnGlyLeuIleThrSerAspTrpGluLeuLeuSerAspProThrProGly	789
Db	2790	AGGGACATTAAACAAGCTGGAGGAGAGAACATCCAGACATTGTTTCAGAA	2837
Oy	790	IleProSerProArgAspArgLysLeuCysValAlaGlyAlaGlnLeuTrpValAcysGlnAsp	809
Db	2838	---AAGCAGCCACMACAGAGTG-----GAC	2861
Oy	810	ProAlaIlePheGlnGluArgHisLeuLysTrpIleSerLeuLeuGlyLysGlyAsnPhe	829
Db	2862	CCCCCTCACTTGAAGACCGGCTTCGAGAGAGATTCCGATCTGGAGAGGGTCACATT	2921
Oy	830	GlySerValGluLeuCysArgTrpAspProLeuGlyAspAspThrGlyProLeuValAla	849
Db	2922	GGGAGAGTTAGCTCTGCACATATGATCTCGAGGAGACACACAGGGGAGCAGGTAAGTCT	2981
Oy	850	ValLysGlnLeuGln---HisSerValProAspGlnGlnArgAspPheGlnArgGluIle	868
Db	2982	GTCACATCCCTGAAGCGCTGAGATGGAGGATACCACTATGCTAGCTCAAGAGAGATA	3041
Oy	869	GlnIleLeuLysAlaLeuHisSerAspPheIleValLysTrpArgGlyValSerTrpGly	888
Db	3042	GAGATCTTACGGAACCTTACATGAGAACTTGTGAATACAAAGAGATTCGATGGAA	3101
Oy	889	ProGlyArgGlnSerLeuArgLeuValMetGluTrpLeuProSerGlyCysLeuArgAsp	908
Db	3102	GACGAGGCAATGGTATCAAGCTCATCATGAGATTCTCGCTGGGAAGCCTAAAGAG	3161
Oy	909	LeuLeuGlnArgHisArgGly---LeuHisThrAspArgLeuLeuPheAlaTrpGln	927
Db	3162	TATTTGCCAAAGATATAGAACAAATAACCTCAACACACACCTTAAATATGCAATCCAG	3221
Oy	928	IleCysLysGlnMetGluTrpLeuGlnGlyAlaArgArgCysValHisArgAspLeuAlaIa	947
Db	3222	ATTTGTAAAGGGGATGCACACTCTGGGTCTCGGCAATACGTTTACACGGGACTTACAGCA	3261
Oy	948	ArgAsnIleLeuValGluSerGlnAlaHisValLysIleAlaAspPheGlyLeuAlaLys	967
Db	3282	AGAAATGCTCTTGTGAGACTGAGCATCAAGAGAGAGATCGGAGACTTGTGTTAAACAA	3341
Oy	968	LeuLeuProLeuGlnLysAspTrpTrpValValAlaValGlnProGlnGlnSerProIlePhe	987
Db	3342	GCAATTAAACCCGATTAAGGAGTACTACACAGTCAAGAGAGACGGGAGACGCCAGGTGTC	3401
Oy	988	TrpTrpAlaProGlnSerLeuSerAspAsnIlePheSerArgGlnSerAspValTrpSer	1007
Db	3402	TGGTAGCCTCCGGATAGTTTAATCCAGTGAATTTTATATCGCCTCGATGCTCGGCT	3461
Oy	1008	PheGlyValValLeuTrpGlnLeuPheThrTrpCysAspLysSerCysSerProSerAla	1027

```

Db      3462 TTGGAGAGTGACACCTCCAGAGCTGCTCACTTACCTAGCTAGACTATTGATCCATGCGC 35
Qy      1028 GTupheLeuArgMetMetGlyPRoGluArgGluGlyPRoLeuGlySArgLeuLeuGlu 104
      |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db      3522 TTGTTCTCGGAATAATGATAGGCGCAACCTCATGCGCCAGATGAGACAGTGACAGCGCTTGAG 358
Qy      1048 LeuLeuAlaGluGlyAArgARLeuPRoPRoPRoPRoPRoPRoPRoPRoPRoPRoPRoPRo 108
      ||| |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db      3582 ACTCTGAAGAAGAGAGAGAGCGTGTGCCATGTCTCCACCAACTCTCTCGATGAGGCTTATCAG 364
Qy      1068 LeuMetGlnLeuGlySTRPalAPRoGluPRoHisAspArgPRoAlaPheAlaThrLeu 1086
      |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db      3642 CTTATGAGAAATAATGCTGGGAATTCACACCATCTACCGGACACACTTTACAGACCTT 3698

RESULT 2
US-09-880-107-2379
?
? Sequence 2379, Application US/09880107
? Patent No. US20020142981A1
? GENERAL INFORMATION:
? APPLICANT: Horne, Darci T.
? APPLICANT: Wockley, Joseph G.
? APPLICANT: Scherf, Uwe
? APPLICANT: Gene Logic, Inc.
? TITLE OF INVENTION: Gene Expression Profiles In Liver Cancer
? FILE REFERENCE: 44921-5028-NO
? CURRENT APPLICATION NUMBER: US/09/880,107
? CURRENT FILING DATE: 2001-06-14
? PRIOR APPLICATION NUMBER: US 60/211,379
? PRIOR FILING DATE: 2000-06-14
? PRIOR APPLICATION NUMBER: US 60/237,054
? PRIOR FILING DATE: 2000-10-02
? NUMBER OF SEQ ID NOS: 3950
? SOFTWARE: PatentIn Ver. 2.1
? SEQ ID NO 2379
? LENGTH: 3541
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
? OTHER INFORMATION: Genbank Accession No. US20020142981A1 M64174
US-09-880-107-2379

Alignment Scores:
Pred. No.: 5 33e-191 Length: 3541
Score: 1904.00 Matches: 443
Percent Similarity: 53.66% Conservative: 188
Best Local Similarity: 37.67% Mismatches: 381
Query Match: 32.49% Indels: 164
DB: 10 Gaps: 24

US-09-397-967-16 (1-1099) x US-09-880-107-2379 (1-3541)
Qy      17 SerLeuSerSerSerGluAlaGlyAlaLeuHisValLeu---LeuPRoPRoArgGlyPRo 35
      ||||| :||: ||| ||| :||| ||| ||| |||
Db      121 AACCTGAGAGCGCCCTGAGCCAGGCGTGGAAAGTATCTTCTATGCTGCGAGACGAGCC 180
Qy      36 GlyPRoPRoGlnArgLeuSerPheSerPheGlyAspArgLeuAlaGluAspLeuGlyVal 55
      ||| |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db      181 -----CTCCGGCTGGAGTGGAGAGTACACAGAGAGGAAGCTGTGCATC 222
Qy      56 ArgAlaAlaLysAlaGlyGlyLeuPRoValTyrHisSerLeuPheAlaLeuAlaThr 75
      |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db      226 AGGCGTGGACAGAGGATGCGGATCTCTCTCTTTGTACACACCTCTTGGCCCTGTATGAC 285
Qy      76 GlnAspPheSerGlySTRPhePRoPRoSerHisIlePheCysIleGluAspValAspThr 95
      |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db      286 GAGAACACCAAGACCTCTGTATGCTCAAAATGCAACATCACACCGTTTATGACAAAGATGTC 343
Qy      96 GlnValLeuValTyrArgLeuArgPheTyrPhePRoAspTRPheGlyLeu----- 112
      ||| |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db      346 CTCGGGCTCCACATACCGGATAGGTTCTATTTTACCAATTTGGATGGACACCAACGACAAT 405
Qy      113 ---GlnThrCysHisIArgPheGlyLeuArgLysAspLeu----- 124
      |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

```

```
Dh 406 GAGACGTACGTGGCTCATTTCTCCAAAGACGAGAAAATGGCTACGAGAAAAAAG 465
Oy 125 -----ThSerAlaIleuAspLeuHisValLeuGluHisLeuPheIleAlaGlnHis 141
Dh 466 ATTCACATGCAACCCCTCTTGATGACGACTGAGATATCTGTTGCTCAGGGA 525
Oy 142 ArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLysGluGln----- 159
Dh 526 CAGATGATTTGGTGAATACCTGGCTCTATTGAGACCCCAAGCCGAGCGAGATGGA 585
Oy 160 -----GlyIlePheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArg 175
Dh 586 CAGATATTGAGAACGAGTGTGATGAGATGCTGCTGCGCATCTCACAATAATGCAATG 645
Oy 176 GluGlnAlaGlnArgProGlyGlyLeuLeuLysThrValSerTyrLysAlaCysLeuPro 195
Dh 646 ATGAGAGATGACGTTGCCAGAACTGCCAAGACATCAGATCAACAGCATATATTTCA 705
Oy 196 ProSerLeuArgAspValIleGlnGlyGlnAsnPheValPheArgArgArgIle----- 213
Dh 706 GAAACATTTGATTAAGTCCATCAGACAGAGACCTTCTCACAAGATGCGGATTAATAT 765
Oy 214 -----ArgArgThrValValLeuAlaLeuLeuPro 223
Dh 766 GTTTTCAGAGATTCTCTAAGAGATTATTAACACAGACCATTTGTGACAGACGGTG--- 822
Oy 224 CysGlyArgLeuProGlyArgProTyrAlaLeuMetAlaLysTyrIleLeuAspLeuGlu 243
Dh 823 -----TCCACGCTACCTCGAAGTGAATATCTTGCTACCTTGGAA 864
Oy 244 ArgLeuHisProAlaIleAlaThrThrGluThrPheArgValGlyLeuProGlyAlaGlnGlu 263
Dh 865 ACCTTGACAAACACTTACGGTGTGGAATATTT----- 897
Oy 264 GluProGlyLeuLeuArgValAlaGlyAspAsnGlyLeuProTyr---SerSerAsnAsp 282
Dh 898 GAGACTTCATGTTACTGATTTTATCAGAAAATGAGATGATGATTGTTTCATTCCGATGAC 957
Oy 282 ----- 282
Dh 958 GGTGGAACGTTCTCTACTACGAAAGTATGCTGCGAATCTTGGAAATCCAGTGGAGG 1017
Oy 282 ----- 282
Dh 1018 CATTAACCAATGTTGTTCTGTGAAAAGAAAAAATAAATCAAGACGGAATAAATCTG 1077
Oy 283 -----GluLeuPheGlnThrPhe 288
Dh 1078 GAAATTAAGACAAAGAGATGAGAGAAAAACAAGATCCGGGAGAGTGGAAACAATTTT 1137
Oy 289 CysAspPheProGluIleValAspValSerIleAsnGlnAlaProArgValGlyProAla 308
Dh 1138 TCATCTTCCTCGAATCAGCACATGTATATAAGAGTCT----- 1179
Oy 309 GlyGluHisArgLeuValThrValThrArgMetAspLysIleIleLeuGluAlaGluPhe 328
Dh 1180 -----GTGCTCAGATTAAACAAGCAGCAACAAGAAAAGAACTCAAGCTC 1227
Oy 329 ProGlyLeuProGluAlaLeuSerPheValAlaLeuValAspGlyTyrPheArgIle 348
Dh 1228 TCTTCCACGAGGAGGCTTGTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1287
Oy 349 CysAspSerArgHisTyrPheCysLysGluValAlaProProAlaGluLeuGluGlu 368
Dh 1288 GCAGATGCCATCATTTACCTCTGACGAGCTGCCCCCGGTGATGCTCCACACATA 1347
Oy 369 AlaAspValCysHisGlyProIleThrLeuAspPheAlaIleHisLysLeuLysAlaAla 388
Dh 1348 CAGAAATGCGTGTCAATGCTCAATCTGTACAGAAATACCCATCAATAATTCCGCAAGAA 1407
Oy 389 GlySerLeuProGlyThrTyrIleLeuArgArgSerProGlnAspTyrAspSerPheLeu 408
Dh 1408 GGAAGCGAGAGGAGATGCTGAGTGTGAGCTGACACGACTTGTGACACATCTCTC 1467

Oy 409 LeuThr---AlaCysValGlnThrPro-----LeuGlyProAspTyrLys 422
Dh 1468 ATGACGCTACCTGCTTTGGAAGTCTGACAGAGTGCAGGGTCCCAAGAACGTTCAAG 1527
Oy 423 GlyCysLeuIleArgGlnAspProSerGlyAlaPheSerLeuValGlyLeuSerGlnPro 442
Dh 1528 AACTTTCAGATC---GAGGTGCAGAAAGGCGGCTACAGTCTGCACGGTTGGACCGGAGC 1584
Oy 443 HisArgSerLeuArgGlyLeuLeuAlaAlaCysTrpAsnSerGlyLeuArgValAspGly 462
Dh 1585 TTCCCGCAGTTGGAGACCTCATGAGCCACCTCAGAAAGCAGATCTCCGACGATAC 1644
Oy 463 AlaAlaLeuTyrLeuThrSerCysAlaProArgProLysGlyLysSerAsnLeuIle 482
Dh 1645 ATCAGCTTCATGCTTAAACCTGTGTCAGCCCAAGGCCGAGAAATCTCCAACTGCTG 1704
Oy 483 ValValArgArgGlyCysAsnProAlaProAlaProGlyCysSerProSerCysAla 502
Dh 1705 GTGGCTACTAAG-----AAGCCAGAGTGGCAGCCC---GCTACCCC 1746
Oy 503 LeuThrGlnLeuSerPheHisThrIleProThrAspSerLeuGluTyrPheIleAsnLeu 522
Dh 1747 ATGAGCCAGCTGAGTTTCATCGGATCTCAAGAGAGATCTGTGACGGCGACACTT 1806
Oy 523 GlyHisGlySerPheThrLysIlePheArgGlySerArgArgGluValAlaAspGlyGlu 542
Dh 1807 GCGAGACGACGAGAACACACATCTATTGAGACCTGTATGATTCAGAGATGAGACGAA 1866
Oy 543 ThrHisAspSerGlu-----ValLeuLeuLysValMetAspSerArgHisArg 558
Dh 1867 GGAACCTCTGAGAGAGAGATTAAGATGATCTCTCAAACTCTTAGACCCCAACAGG 1926
Oy 559 AsnCysMetGluSerPheLeuGluAlaAlaSerLeuMetSerGlnValSerTyrProHis 578
Dh 1927 GATATTTCCCTGGCTCTTTCGAGGACGCGACATGTATGAGACAGGCTCTCCCAAAAC 1986
Oy 579 LeuValLeuLeuHisGlyValAlaCysMetAlaGly---AspSerIleMetValGlnGluPhe 597
Dh 1987 ATGCTGATCCTATAGCGTGTGCTGTCGCGAGTGGAGATATCATGCTGGAAGATTT 2046
Oy 598 ValTyrLeuGlyAlaIleAspMetTyrLeuArgLysArgGlyHisLeuValSerAlaSer 617
Dh 2047 GTGGAAGGGGCTCTCTGATCTCTTCATGACCCGGAAGAGATGCTCTTACCACACA 2106
Oy 618 TrpLysLeuGlnValThrLysGlnLeuAlaTyrAlaLeuAsnTyrLeuGlnAspLysGly 637
Dh 2107 TCGAAATTCAAAGTGGCCAAACAGCTGGCCAGTCCCTGAGCTACTTGAGGATTAAGAC 2166
Oy 638 LeuProHisGlyAsnValSerAlaArgLysValLeuLeuAlaArgGlyGlyAspGly 657
Dh 2167 CTGGTCATGGAATGTGTACTAATAAAACCTCTCTGCGCCGCGTGGGGAATCGACAGT 2226
Oy 658 Asn---ProPheIleIleLysLeuSerAspProGlyValSerProThrValLeuSerLeu 676
Dh 2227 GAGTGTGCCCCATTCATCAAGCTAGTGAACCCGCGATCCCATTTACGCTGCTCTAGG 2286
Oy 677 GluMetLeuThrAspArgIleProTyrValAlaProGlyCysLeuGlnGluAlaGlnThr 696
Dh 2287 CAACAATTCATGAAAGAAATCCCATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2346
Oy 697 LeuCysLeuGluAlaAspLysTrpGlyPheGlyAlaAlaThrThrTrpGluValPheGlnArg 716
Dh 2347 CTGAGTGTGCTGCTGCAAGGTGAGCTTGAACCAACGCTGCGGAAATCTGTACAAAT 2406
Oy 717 GlyProAlaHisIleThrSerLeuGluProAlaLysLysLeuLysPheTyrGluAspGln 736
Dh 2407 GGCAGAGATCCCTTGAAGAACACAGACGCTGATTAAGAGAAAGAAATCTTAAGAACCGG 2466
Oy 737 GlyGlnLeuProAlaLeuLysTrpThrGluLeuAlaGlyLeuIleThrGlnCysMetAla 756
Dh 2467 TGCAGGCGAGTGCACCATCATGTAAGAGCTGCTACCTCATGACCCGCTGCTGATAC 2526
```



```
Db 765 GTTTCAGAGATTTCCTTAAGAAATTTACACAGACCATTTTGTGACGACGCGTG--- 821
Oy 224 CysGlyArgLeuProGlyArgProTyrAlaLeuMetAlaTyrTyrIleLeuAspLeuGlu 243
Db 822 -----TTCACGCATGCCTGAAGGTGAAGTAATCTGTGCTACTCTGGAA 863
Oy 244 ArgLeuHisProAlaIleThrThrGluThrPheArgValGlyLeuProGlyAlaGlnGlu 263
Db 864 ACTTTGACAAAACATTACGCTGCTGAATATTT----- 896
Oy 264 GluProGlyLeuLeuArgValAlaGlyAspAsnGlyIleProTyr---SerSerAsnAsp 282
Db 897 GAGACTTCATGTTACGTACATTTTCATCAGAAATGATGATGATGTTGTTTCATTCGATAC 956
Oy 282 ----- 282
Db 957 GGTGAAAACGTTCTCTACTACGAAGTGTGATGCGGAATCTTGGAAATCCAGTGAGG 1016
Oy 282 ----- 282
Db 1017 CATAAACCAATGTTGTTTCTGTTGAAAAAGAAAAATAACTGAAACGGAAGAAAAATCG 1076
Oy 283 -----GluLeuPheGlnThrPhe 288
Db 1077 GAAATATAACACAGAGATGAGAGAAAAACAAGATCCGGAAAGAGTGGAACAAATTTT 1136
Oy 289 CysAspPheProGluIleValAspValSerIleAsnGlnAlaProArgValGlyProAla 308
Db 1137 TCTTACTCCCTGGAATATCACACATTTGTAATAAAGAGTCT----- 1178
Oy 309 GlyIleHisArgLeuValThrValThrArgMetAspGlyHisIleLeuGluAlaGlnPhe 328
Db 1179 -----GtGGTACGCTTAACAGACAGACAGACAAAGAAATGGAACGTAGCTC 1226
Oy 329 ProGlyLeuProGluAlaLeuSerPheValAlaLeuValAspGlyTyrPheArgLeuIle 348
Db 1227 TCTTCCACGAGAGGCGCTTCTGCTTGTGCTCCGTGAGATGCGCTACTTCGCGCTACA 1286
Oy 349 CysAspSerArgHisTyrPheCysGlyValAlaProProArgLeuLeuGlnGluIle 368
Db 1287 GCAGATGCCCATTCATCTCTGCACGCGGCCGCCCGTGTGATGCTCCACACACTA 1346
Oy 369 AlaAspValCysHisGlyProIleThrLeuAspPheAlaIleHisIleLeuIleAla 388
Db 1347 CAGATGCGCTGTCATGCTGCATCTGTACAGATATCGCCATCAATAATTTGGGCAAGAA 1406
Oy 389 GlySerLeuProGlyThrTyrIleLeuArgArgSerProGlnAspTyrAspSerPheLeu 408
Db 1407 GGAAGCGAGAGGAGTGTAGCTGTGAGTGGGCTGC-ACCGACTTTGACAAACATCTC 1465
Oy 409 LeuThr---AlaCysValGlnThrPro-----LeuGlyProAspTyrIle 422
Db 1466 ATGACCGTCACCTGCTTTGAGAAAGTCTGACGACGTCGAGGGGCCAGAGAGCTTCAG 1525
Oy 423 GlyCysLeuIleArgGlnAspProSerGlyAlaPheSerLeuValGlyLeuSerGlnPro 442
Db 1526 AACTTCAGATC---GAGGTGCAGAGAGGCGCTACAGTCTGCAGGTTGGACCGCAGC 1582
Oy 443 HisArgSerLeuArgGlnLeuLeuAlaIleCysTrpAsnSerGlyLeuArgValAspGly 462
Db 1583 TTCCCGACCTTGTGGAGACCTCATGAGCCACTCAAGAAAGACAGATCTCGCGCATAC 1642
Oy 463 AlaAlaLeuTyrLeuThrSerCysAlaProArgProIleGlySerAsnLeuIle 482
Db 1643 ATCAGCTTCATGCTAAACCGCTGCTGCCAGCCAGCCCGGAAATCTCCAACTGCTG 1702
Oy 483 ValValArgArgGlyCysAsnProAlaProAlaProGlyCysSerProSerCysAla 502
Db 1703 GTGGCTACTAAG-----AAAGCCGAGAGTGGCAGCCC---GTCATACCC 1744
Oy 503 LeuTrgInLeuSerPheHisThrIleProThrAspSerLeuGluTyrHisIleAsnLeu 522
Db ----- 522
Db 1745 ATGAGCCAGCTGATTTGCATCGAGTCCTCAAGAAAGATCTGTGTCAGGCGACACCTT 1804
Oy 523 GlyHisGlySerPheThrLysIlePheArgGlySerArgArgValAlaAspGlyGlu 542
Db 1805 GGGAGAGCGACAGAGAACACACATCTATTCGGGACCTGTAGTGTCAAGAGATGAGAA 1864
Oy 543 ThrHisAspSerGlu-----ValLeuLeuLysValMetAspSerArgHisArg 558
Db 1865 GGAACCTCTGAGAGCAAGAGATAAAGTGCATCTCAAAAGCTTAGACCCCGACACAGC 1924
Oy 559 AsnCysMetGluSerPheLeuGluAlaIleSerLeuMetSerGlnValSerTyrProHis 578
Db 1925 GATATTTCCCTGGCTCTTCAGGACAGCCAGATGATGAGACAGCTCTCCCAACAAAC 1984
Oy 579 LeuValLeuLeuHisGlyValCysMetAlaGly---AspSerIleMetValGlnGluPhe 597
Db 1985 ATCGTGTACTCTATGCGCTGTGTGCGGACGCTGAGATATCATATGATGAGAGATTT 2044
Oy 598 ValTyrLeuGluAlaIleAspMetTyrLeuArgLysArgLysIleValIleValSer 617
Db 2045 GTGGAAGGGGCTCCTCGATCTCTCATGACCGGAAAGCGATGTCCTTACACACCA 2104
Oy 618 TrpLysLeuGlnValThrLysGlnLeuAlaTyrAlaLeuAsnTyrLeuGluAspLysGly 637
Db 2105 TGAATATCAAAAGTGGCCAAACAGCTGCGCCCTGACCTCTGAGAGATAAAGAC 2164
Oy 638 LeuProHisGlyAsnValSerAlaArgLysValLeuLeuAlaArgGlnGlyLysAspGly 657
Db 2165 CTGGTCCATGGAATGTGTGACTAAAACCTCTCTCGCGCCGTGAGGCGATGACAGT 2224
Oy 658 Asn---ProProPheIleLysLeuSerAspProGlyValSerProThrValLeuSerLeu 676
Db 2225 GAGTGTGCGCCCTTCATCAACACTCATGACCGCCGATCCCATACGCTGCTCTAGG 2284
Oy 677 GluMetLeuThrAspArgIleProTyrValAlaProGluCysLeuGlnGluAlaGlnThr 696
Db 2285 CAAGATTCATTTGAACGATCCATGATGCTCTGAGTGTGTGGAGACTCCAGAAC 2344
Oy 697 LeuCysLeuGluAlaAspLysTyrGlyPheGlyAlaThrThrTrgIleValPheGlnArg 716
Db 2345 CTGACTGTGCGCTGACCAATGAGAGCTTGGAACACAGCTTGGAAATCTGTACAT 2404
Oy 717 GlyProAlaHisIleThrSerLeuGluProAlaLysLysLeuLysPheTyrGlnAspGln 736
Db 2405 GCGCAGATCCCTTGAAGAACAGACGCTGATGAGAAAGAGATCTTAAAGAACGCG 2464
Oy 737 GlyGlnLeuProAlaLeuLysTyrThrGlnLeuAlaGlyLeuIleThrGlnCysMetAla 756
Db 2465 TGCAGGCCAGTGACACCATCTGTAAGAGCTGCTGACCTCATGACCCGCTGCATGAC 2524
Oy 757 TyrAspProGlyArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIle 776
Db 2525 TATGACCCCATTCAGAGCGCTTCTTCGAGCCATCATGAGACATTAATAAATCTGAA 2584
Oy 777 ThrSerAspTyrGlnLeuLeuSerAspProThrProGlyIleProSerProArgAspGlu 796
Db 2585 GAGCAGAAATCCAGATATTGTTTCAGAAAAAACCACCACTGAA----- 2629
Oy 797 LeuCysValAlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGlnGluArg 816
Db 2630 -----GTGGACCCACACATTTTGAAGAGCGC 2656
Oy 817 HisLeuLysTyrIleSerLeuLeuGlyLysGlnAsnPheGlySerValGluLeuCysArg 836
Db 2657 TTCTTAAGAGAGATCCGAGCTTGGAGAGGCGCACTTTGGAGAGTGTGACCTCTGCGAG 2716
Oy 837 TyrAspProLeuGlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGln---His 855
Db 2717 TATGACCCCGAAGGAGCAATACAGGAGGAGGAGCTGTAAATCTCTAAGCTGTAG 2776
Oy 856 SerValProAspGlnIleArgAspPheGlnArgGlnIleGlnIleLeuLysAlaLeuHis 875
Db 2777 AGTGAAGGTAAACCATGATGCTGATCTGAAAAAGAAATCGAGATCTTAAAGAACCTCTAT 2836
```


Qy 483 ValValAlrgratgGlyCysAsn-ProAlaProAlaProGlyCysSerProSerCys-CysA 502
 Db 1917 -----CGAGGTGCTCTCCTGTCGACAGCAGCAGCAGTATCCATCATGATATGTG 1967
 Qy 502 JaleuThInLseuSerPheHisThrIleProThrAsp 514
 Db 1968 CTGAGACCCAGATGACCAAAACAGCGTG--ACTGAC 2002

RESULT 6

US-09-962-854A-5
 : Sequence 5, Application US/09962854A
 : Publication No. US20030044803A1
 : GENERAL INFORMATION:
 : APPLICANT: Pedersen, Finn S.
 : APPLICANT: Annette, Soerensen B.
 : APPLICANT: Hernandez, Javier Martin
 : TITLE OF INVENTION: Methods for diagnosis and treatment of diseases associated with a
 : TITLE OF INVENTION: expression of JAK1
 : FILE REFERENCE: A-70020/RMS/DCF
 : CURRENT APPLICATION NUMBER: US/09/962,854A
 : CURRENT FILING DATE: 2001-09-24
 : PRIOR APPLICATION NUMBER: US 09/668,644
 : PRIOR FILING DATE: 2000-09-22
 : NUMBER OF SEQ ID NOS: 5
 : SOFTWARE: Patent version 3.1
 : SEQ ID NO 5
 : LENGTH: 975
 : TYPE: DNA
 : ORGANISM: Mus musculus
 : FEATURE:
 : NAME/KEY: misc_feature
 : LOCATION: (799)..(799)
 : OTHER INFORMATION: "n" at position 799 can be any base.
 : US-09-962-854A-5

Alignment Scores:

Pred. No.: 1,51e-41 Length: 975
 Score: 483.50 Matches: 131
 Percent Similarity: 61.62% Conservative: 52
 Best Local Similarity: 44.11% Mismatches: 94
 Query Match: 8.25% Indels: 23
 DB: Gaps: 9

US-09-397-967-16 (1-1099) x US-09-962-854A-5 (1-975)

Qy 813 PheGluGluArGHisLeuLysTyrlIeSerLeuLeuGlyLysGlyAsnPheGlySerVal 832
 Db 8 TTTGAAAAGCGTTCCTGGAAGAGATTCGTGACTTGGAGAGGTCACCTTGGCAAGGTT 67
 Qy 833 GluLeuGlyArGlyrAspProLeuGlyAspAsnThrGlyProLeuValAlaValLysGln 852
 Db 68 GAGCTCTGACATATGATTCCTGAGGAGACAAACAGGAGGAGCGATGCTCAAGTCC 127
 Qy 853 LeuGln---HisSerValProAspGlnGlnArGAspPheGlnArGlyLysGlnIleLeu 871
 Db 128 CTGAAGCGTGAAGAGTGAACACATAGCTATCGAAGAGGATGAGATCTTA 187
 Qy 872 LysAlaLeuHisSerAspPheIleValLysTyrlArGlyValLysTyrlGlyProGlyVal 891
 Db 188 CGGAACCTCTACATGAGAACATTTGTGAGTACAAAGGAATCTGCATGGAAGAGGAGGC 247
 Qy 892 GlnSerLeuArGLeuValMetGlyLysLeuProSerGlyCysLeuArGAspLeuLeuGln 911
 Db 248 AATGCTATCAACGTCATCATGAGTTCCTCGGAAACCTTAAGAGATATCTGCCA 307
 Qy 912 ArgHisArGly---LeuHisThrAspArgLeuLeu-LeuPheAlaTrpGlnIleCysGly 930
 Db 308 AAGAAATAGAAACAACACCTCAACACAGCTAAATAAATATGCCATCCAGAAATTGTA 367
 Qy 930 sGlyMetGlyLysLeuGlyAlaArGArGValHisArGAspLeuAlaAlaArGAsnI 950
 Db 368 GGGATGACACTTGGCTTCTCGGCAATAGTTCACCGGAGCTTAGCAGCCAG-AAATGT 426

Qy 950 eLeuValGlnSerGluAlaHisValLysIleAlaAspPheGlyLeuAlaLysLeuLeuP 970
 Db 427 CCTTGTGAGATGGATCCAGTTCAGATTCGAGACCTTGAGTTAACCAAGCATTTG 486

Qy 970 oLeuGlyLysAspTyrlValValArGLeuProGlyLysLeuProIlePheTrpVal 990
 Db 487 AAAGATTAGAGTACACACAGTTCAGACACCGGAAAGACAGTTCGCGTACGC 546

Qy 990 aProGlySerLeuSerAspAsnIlePheSerArGlnSerAspValTrpSerPheGlyVal 1010
 Db 547 TCCGGAATGTTTAAATCCAGTGTATTTTAAACGCCCTCCCATGTCCGCTTGGAGT 606

Qy 1010 lValLeuTyrlGlnLeuPheThrTyrlCysAspLysSerCysSerProSerAlaGluPheLe 1030
 Db 607 GACACTGCAGCAGCTGCTCATTTACTGTGACTCCGAATTTAGTCCATGGCCTT----- 660

Qy 1030 uArgHecMetGlyProGluArGlyGluPro---ProLeuGlyArGLeuLeuGln---- 1047
 Db 661 -----GTCGCCGAAAGGTAAAGCCCAACTCCAGCCAGAAACAAATTGAAGGCT 710

Qy 1048 -LeuLeuAlaGluGlyArGArG---LeuProProProProThrCysProThrGluValG 1066
 Db 711 GTGCATCTACTAAGAAAGAGAGAGCCCTGCGATGTCACCCAA-TGTCCTGATGAAGTTAA 769

Qy 1066 nGluLeuMetGlnLeuGlyTrpAlaProGlu-----ProHisAsp---ArgProAlaP 1083
 Db 770 CAGCTTATGGAGAAATTCCTGGAATTCGACTACTAACCAACAATTTCCGAACCTATG 829

Qy 1083 eAlaThrLeuSerProGlnLeuAspProLeuTrpArgLysArgProGly 1099
 Db 830 GAAGAGTTTAAAGCCCC-----TTTAAATAGAAAGCTGCG 863

RESULT 7

US-09-910-943-636
 : Sequence 636, Application US/09910943
 : Patent No. US20020081610A1
 : GENERAL INFORMATION:
 : APPLICANT: Hemmati-Brianlou, Ali
 : APPLICANT: Altman, Curtis
 : TITLE OF INVENTION: Assays and Materials for Embryonic Gene Expression
 : FILE REFERENCE: 7529/JG1480S1
 : CURRENT APPLICATION NUMBER: US/09/910,943
 : CURRENT FILING DATE: 2001-07-23
 : NUMBER OF SEQ ID NOS: 742
 : SOFTWARE: Patent version 3.1
 : SEQ ID NO 636
 : LENGTH: 753
 : TYPE: DNA
 : ORGANISM: Xenopus laevis
 : FEATURE:
 : NAME/KEY: misc_feature
 : LOCATION: (1)..(763)
 : OTHER INFORMATION: n may be a or g or c or t/u
 : US-09-910-943-636

Alignment Scores:
 Pred. No.: 1.31e-41 Length: 753
 Score: 482.50 Matches: 102
 Percent Similarity: 60.96% Conservative: 37
 Best Local Similarity: 44.74% Mismatches: 71
 Query Match: 8.23% Indels: 18
 DB: Gaps: 2

US-09-397-967-16 (1-1099) x US-09-910-943-636 (1-753)

Qy 660 ProPheIleLysLeuSerAspProGlyValSerProThrValLeuSerLeuGluMetLeu 679
 Db 76 CCATTCAATAAGCTTAGATCGATCGAATACCAATTACTGTGTTAAACAACAAGACGG 135

Qy 680 ThrAspArgIleProTrpValAlaProGlyLysLeuGlnValGlnThrLeuGlyLeu 699
 Db 136 GTAGAGCCATTTCATGATGATGCTCTGAAATGTGTTAGAGATTCCAGAGATTAAAGTGT 195


```

OY 700 GUAAlaAspLysTrpGlyPheGlyAlaThrThrTrpGluValPheGlnArgGlyProAla 719
    |||||||  |||||  |||  |||||||
Db 196 GCGCGGACAAATGAGACTTTGGACCACTATGGGAATCTGTATCAATGGAGAAGTG 255
OY 720 HisIleThrSerLeuGluProAlaLysLysLeuLysPheTyrGluSpGlnGlyGlnLeu 739
    ::|  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 256 CCRCTCAAAAGCCAAAGCCCTGACAGAGAAAGAAAGATTCTACGAGAGATCTTCATCTTA 315
OY 740 ProAlaLeuLysTrpThrGluLeuAlaGlyLeuIleThrGlnCysMetAlaTyrAspPro 759
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 316 GTGGCACCTTCATCCAAAGAGTTAGCAGATCTATATATCATTCATGATATTCAGACCC 375
OY 760 GlyArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIleThrSerAsp 779
    |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
Db 376 AACGAGAAGACCGTTTATTAGCAATCATGAGAAATCAACAAGCTGGAAGAGCAAAAT 435
OY 780 TyrGluLeuLeuSerAspProThrProGlyIleProSerProArgAspGluLeuGlyVal 799
    ::|  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 436 CCAGACATTCCTCTGTAACAAAACACCATCT- 465
OY 800 AlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGlnGluArgHisLeuLys 819
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 466 -----GCGGAAGTGGATCCAACTTATTGGAGAAATATTCCTTCAAG 507
OY 820 TyrIleSerLeuLeuGlyLysGlyAsnPheGlySerValGluLeuGlyTyrAspPro 839
    ::|  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 508 AGAGTAAGAGATCTAGAGAGAGGCGCATTTTGGAAAGTTGAATATGATAGTATGACCCA 567
OY 840 LeuGlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGlnHisSerValProAsp 859
    |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
Db 568 GAGGCGACAAACAGGGGAGACTGGTGTATTAATCGCTAAAGCCCTGGCACAGGGGCG 627
OY 860 GlnGlnL-ArgAspPheGlnArgGluIle-GlnIleLeuLysAlaLeuHisSerAsp 878
    ::|  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 628 AGCGCATTCCTGATCTGAANAANGAATTCGAATCTCAGAAATCTGATCATGAGAA 687
OY 878 eIleValLysTyrArgGlyVal 885
    |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
Db 688 TATTGTCAATATCAAGAAGATT 709

RESULT 8
US-10-044-090-148
; Sequence 148, Application US/10044090
; Patent No. US20020137081A1
; GENERAL INFORMATION:
; APPLICANT: Olga Bandman
; TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
; FILE REFERENCE: PA-0028 US
; CURRENT APPLICATION NUMBER: US/10/044, 090
; CURRENT FILING DATE: 2002-01-09
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: PERL Program
; SEQ ID NO. 148
; LENGTH: 3914
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20020137081A1 3331519CB1
US-10-044-090-148

Alignment Scores:
Pred. No.: 1,75e-40 Length: 3914
Score: 482.00 Matches: 311
Percent Similarity: 35.45% Conservative: 152
Best Local Similarity: 23.81% Mismatches: 402
Query Match: 8,23% Indels: 447
DB: 12 Gaps: 66

US-09-397-967-16 (1-1099) x US-10-044-090-148 (1-3914)
OY 12 ProGlnArg-----SerCysSerLeuSerSerSerLeuAlaGlyAlaLeuHisVal 28
    |||||||  |||  |||||||

```

```

Db 203 CCCCAGCGCTCTTCTCTGACTTGC-----GTGCTGGGAGGCGCGGCGG----- 247
OY 29 LeuLeuProProArgGly----- 34
Db 248 -----GGGAGGGCTTGGACGCTGGGGGCGCCCTGCTGCTGGAGAAAGAGAC 298
OY 35 -----ProGlyProProGlnArgLeuSer----- 42
Db 299 CGTATCGTGGCACACCCCGCGGCGCACCCCTGCGCTGGCGCAACGGTTGCGACAG 358
OY 43 -----PheSerPhe-----GlyAspTyrLeuAlaGluAspLeuGlyVal 55
Db 359 GTCACGCTTGGCGGCTTCTTCAAGCCCTCGGACCTCGTGGCGCTTCTTCTGCTGGCG 418
OY 56 ---ArgAlaLysAlaCysGlyIleLeuProValTyrHisSerLeuPheAlaLeuAla 74
Db 419 GGTGCTGGGCGGCGCGCACCGCGCTATCTAGCTGCACAACAGC----- 463
OY 75 ThrGluAspPheSerCysTrpPheProProSerHisIlePhe----- 88
Db 464 -----CTGAGAGCCACCTCTCTCCAGACAGTCCACAC 499
OY 89 CysIleGluAspValAspThrGlnValLeuValTyrArgLeuArg----- 103
Db 500 ACTGTACACAAGGTACACCGCTGACTTCTGACGCTGTGCACAGACAGAAAGACACAG 559
OY 104 -----PheTyrPheProAspTrpPheGlyLeuGlu 113
Db 560 GACGTGATCTGGAAGAGCAACGAGATCTTACTTACACCCCTGGACGTGGCATGAAGCCAG 619
OY 114 ThrCysHisArgPheGlyLeuArgLysAspLeuThrSerAlaIleLeuAspLeuHisVal 133
Db 620 GAT---GGGCGGTTCCTGCTGACAGCT---CCCAATGTCACACCCATCGACGGGCACTGA 675
OY 134 LeuGlnHis-----LeuPheAlaGlnHisArgSer 143
Db 676 CAGTCCCATCTTACCTGAGACCCACCCCTGGCGACGCGCTTCTTCTGCGCTACATCGTGG 735
OY 144 AspLeuValSerGlyArgLeuProValGlyLeuSerMetLysGlnGlnGlyIlePheLeu 163
Db 736 GG-GTT-----GTGGGCTGGCGCGCTGGGCGGCGAG----- 764
OY 164 SerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGlnIleAlaGlnArgProGlyGlu 183
Db 765 -----CCTTACCAAGAGAGTGGCCAGGTTGCTTACATGAGAGTGTTGCGACAGCATG 818
OY 184 LeuLeuLysThrValSerTyrLysAlaCysLeuProProSerLeu----- 198
Db 819 ACGGCGAATGTATATGCCCCCTGGCTTCACTGGACCCGCTGTGAACAGCGCTGCAGAG 878
OY 199 ArgAspValIleGlnGlnGlnAsnPheValThrArgArgArgIleArgArgThrValVal 218
Db 879 AGGCGCGTTTGG-----GGCAGAGCTCGCAGAGCAGACAGTGGCC 914
OY 219 LeuAlaLeuLeuProCysGlyArgLeuProGlyArgProTyrArgIleLeuMetAlaLysTyr 238
Db 915 CAGGCATATCAGGCTGCGGGGCGCTCACTTGTGCTGCCACAGACCCCTATGAGTGGCT 971
OY 239 IleLeuAspLeu-----GluArgLeuHisPro----- 247
Db 972 CTGTGTGATCTGGCTGGAGAGAGCAAGCCAGTTCGCGCCCTGGTCATTTTG 1031
OY 248 -----AlaAlaThrThrGluThrPhe 254
Db 1032 GGGCTGATTCGCCACTCCAGTCCAGTGCAGATGTGATGGTATGACTTGGAGCGTTGAGTG 1091
OY 255 ArgValGlyLeuPro---GlyAlaGlnGlnGluProGlyLeuLeuArgValAlaGlyAsp 273
Db 1092 GTTGTGTCTGCCCTCTGGTGGCAGTGGAG-----TGCAGCTGTGAAGAGTCAAG 1139
OY 274 AsnGlyIleProTrpPheSerSerAsnAspGluLeuPheGlnThrPheCysAspPheProGlu 293
Db 1140 ACCGGATCCCGCCAGATCTCAACA-----TGGCCCTCAG 1172

```


QY 294 ILeValAspValSerIleAsnGlnAlaProArgVal-----GlyPro 307
 Db 1173 AA-STGGAGATTCAACTAGACGATGCCCCGGATCACTGACAGCTGCAGGAACCC 1231
 QY 308 AlaGlyIuHnIAsrGluValThrArgMetAspGlyHisIleLeu----- 324
 Db 1232 TTCCTGGCGGGGAGCATAGAGTACGACCAAGCGACAGTGTCTGTCCAC 1291
 QY 325 -----GluAlaGluPheProGlyLeuProGluAlaLeu 335
 Db 1232 AAGCCATTGTGGACCCAGAGAACACACAGCTGATTC-----GAGTGCCC 1339
 QY 336 SerPheAlaIleValAspGlyTyrPheArgLeuIleCys----- 349
 Db 1340 CGCTTGCTCTGGGACAGTGGTCTTGAG-----TGGCTGTCTCCACATCTGGC 1393
 QY 350 -----AspSerArgHisTyrPheCysIys-----GluValAlaPro 361
 Db 1394 GGCACAGACAGCCGGCTTCAAGGTCAATGTGAAGTCCCGCCGCTGGCTGCA 1453
 QY 362 ProArgLeuLeuGluGluGluAla-----AspValCysHisGlyProIleThrLeuAspPhe 380
 Db 1454 CCTCGCTCTGTACCAAGCAGAGCCGACGTTGTGTCTCCCGCTG----- 1501
 QY 381 AlaIleHisIleLeuIysAlaIleAspGlySerLeuPro-----GlyThrTyrIleLeuArgArg 399
 Db 1502 -----GTCTCGTCTCTGGGAGTGGAGTCCATCTCCACTGTCCGCTGCACCTAC 1549
 QY 400 SerProGlnAspTyrAspSerPheLeuLeuThrAlaCys----- 412
 Db 1550 CGCCCGACAGTACCATCTGGA-CTGGTCGACCATTTGGTGGAGCCCGCAGAACCT 1608
 QY 413 ---ValGlnThrProLeuGlyProAspTyrIysGlyCysLeuIleArgGlnAspProSer 431
 Db 1609 GAGGTAAATGAACCTGAGGCCAAGACAGATACAGTGTCTGTGTGACCTGAGCGGCC 1668
 QY 432 -----GlyAlaPheSerLeuValGlyLeuSerGlnProHisArgSerLeu 446
 Db 1669 ACGGAGAGAGAGAGGGGCG-----CTGGGGGCGCTCCACCCCTCAT---GACCAAC 1716
 QY 447 ArgGluLeuLeuAlaAlaCysTrpAsnSerGlyLeuArgValAspGlyAlaAlaLeuTyr 466
 Db 1717 AGACT-GTCTGAGCGCTTGTGTGACAGCCGTGTGGAGGCGT---GGCATGTGGAGGCA 1772
 QY 467 LeuThrSerCysAlaIleProArgProIysGluIysSerAsnLeuIleValAlaArgArg 486
 Db 1773 CTGACC----- 1778
 QY 487 GlyCysAsnProAla---ProAlaProGlyCys-----SerProSer 499
 Db 1779 GGCTGGAGTGAAGCTGCTGTGCTGTGGCGGCGGCACTGCTGGGCGACGGTTTCC 1838
 QY 500 CysCysAlaLeuThrGlnLeuSerPheHisThrIleProThrAspSerLeuGluTrpHis 519
 Db 1839 TCGTGGCGCTGT----- 1850
 QY 520 GluAsnLeuGlyHisGlySerPheThrIysIlePheArgGlySerArgGluVal 539
 Db 1851 ---GGGACGGGACACGGGGCAG-----GAGCGGGGGGGAACGTC 1888
 QY 540 AspGlyGluThrHisAspSerGluValIleLeuIysValMetAspSerArgHisArgAsn 559
 Db 1889 TCATCCCGCCAGGCGCGCTGCTGTGAGCGGACTCAGCGCTGCACCCACTACAG 1948
 QY 560 CysMetGluSerPheLeuGluAlaIleAspLeuMetSerGlnValSerTyrPro---His 578
 Db 1949 ---CTGGAGTGCAGCTTACACATGCACTGCTGGCGCGCGCTGCGCCCTGCACAG 2005
 QY 579 LeuValLeu-----LeuHisGlyValCysMetAla 588
 Db 2006 GTGCTTCTGCCCCCAGTGGGCTCCAGGCCCGCCGACACCTCCAGCCCAAGGCGCTTCTCA 2065

QY 589 GlyAspSerIleMetValGlnGluPheValTyrLeuGlyAlaIleAspMetTyrLeuArg 608
 Db 2066 GACTCCGAAATTCACAGT----- 2083
 QY 609 LysArgGlyHisLeuValSerAlaSerTrpLysLeuGlnValThrLysGlnLeuAlaTyr 628
 Db 2084 -----ACATGAGAG----- 2092
 QY 629 AlaLeuAsnTyrLeuGlnAspLysGlyLeuProHisGlyAsnValSerAlaArgLysVal 648
 Db 2093 -----CACCGGAGGCTGTGCT---GGCCCAATATCCAAAGTACGTTGTG 2134
 QY 649 LeuLeuAlaArgGlyGlyIysAspGlyAsnProProPheIleLysLeuSerAspProGly 668
 Db 2135 GAGGTCCAGCTGGCTGGGGGTGCAGAGAACCCACTGTGATGACCTGGACAGCGCTGAG 2194
 QY 669 ValSerProThrValLeu----- 674
 Db 2195 GAGACAGACCACTATCCGTGGCTCAACGCCAGCAGCGCTACCTCTTCCGCATGCCG 2254
 QY 675 ---SerLeuGluMetLeuThrAspArgIleProTrpValAlaProGlyCysLeuGlnIu 693
 Db 2255 GCCAGCATTCAGGGCTCGGGGAC-----TGG----- 2281
 QY 694 AlaGlnThrLeuCysLeuGluAlaAspLysTrpGlyPheGlyAlaThrThrTrpGluVal 713
 Db 2282 ---ACCAACACAGTAAAGATGCCACCCGCGCAAGCGGCTGCAGGCTGAGGGCCCA 2335
 QY 714 PheGlnArgGlyProAlaHisIleThrSerLeuGluProAlaLysLysLeuLysPheTyr 733
 Db 2336 GTTCCAAAGAGAGCGCGGCTGTGAAGAGGCTGTGATCAGCAGCTGATCTCCGGCTGTG 2395
 QY 734 GluAspGlnGlyGluLeuProAlaLeuLysTrpThrGluLeuAlaGlyLeuIleThr 752
 Db 2396 ---GCTCGCTGTCTGCCACCTGCTCACCATCTGCGCCCTTTTAAACCTG 2446
 QY 753 -----GlnCysMetAlaTyrAspProGlyArgArgProSerPhe 765
 Db 2447 GTGTGATCCGCGAGAGCTGCTG-----CATCGGAGCGACCTTCACCTAC 2494
 QY 766 ---ArgAlaIleLeuAlaArgAspLeuAsnGlyLeuIleThrSerAsp 779
 Db 2495 CAGTCAGGCTCGGCGGAGAGACCATCTGCACTTGCAGTGCAGGAGCTTGACACTTACC 2554
 QY 780 TyrGluLeuLeuSerAspProThrProGlyIleProSerProArgAspGluLeuCysVal 799
 Db 2555 CGCGGCGCAAACTGCAGCGCCGAGCC---CTGAGCTACCA----- 2593
 QY 800 AlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGluGlnArgHisLeuLys 819
 Db 2594 -----GTCTAGAGTGGGAGGACATCCAC 2617
 QY 820 TyrIleSerLeuLeuGlyLysGlnAspPheGlySerVal-----GluLeuCysArgTyr 837
 Db 2618 TTTGAGAGCTCTCTCGGGAGGGAATCTGGGAGGTCACTCCGGGCGCATGACAAGAG 2677
 QY 838 AspProLeuGlyAspAsnThrGlyProLeuValAlaValLysGlnLeu---GlnHisSer 856
 Db 2678 GACGGGCTGAAGTGAAC-----GCAAGCATCAAAATGCTGAAGAGGTATGCC 2725
 QY 857 ValProAspGlnGlnArgAspPheGlnArgGluIleGlnIleLeu---LysAlaLeuHis 875
 Db 2726 TCTGAATAATGACCATGTGACTTTGGCGGAGAACTGGAAGTTCTGTGCAAAATGGGGCAT 2785
 QY 876 SerAspPheIleValLysTyrArgGlyValAspTyrGlyProGluArgGlnSerLeuArg 895
 Db 2786 CACCCCAACATCTCAACCTCTGGGGGCTGT---AAGAACGAGGTTACTTGTAT 2839
 QY 896 LeuValMetCysTyrLeuProSerGlyCysLeuArgAspLeuLeuGlnArgHisIleArgGly 915
 Db 2840 ATCGCATATGAATATGCCCCCTACGGGAAACCTGTATTTTCTGGGAAACCCGGGTC 2899
 QY 916 LeuHisThrAsp----- 919

```

Db 2900 CTAGAGACTGACCCAGCTTTTGTCTGAGAGCATGGAGACGCTCTACCTTAGCTCCGG 2959
      ||| |||||
Qy 920 ArgLeuLeuPheAlaIlePheGlnIleCysIleGlyMetGlyIleValAlaArgArg 939
      :|||:|||||
Db 2960 CAGCTGTGGCTTTGGCAGTGAATGCGCCCAATGCGCATGAGTCTGAGTGAAGAGAG 3019
      :|||:|||||
Qy 940 CysValHisArgAspLeuAlaAlaArgAsnIleLeuValGlySerGluAlaHisValLys 959
      :|||:|||||
Db 3020 TTGATCCACAGAGGACCTGGCTGCGCGGAAATGCTGCTCGAGAGAACTAGCTCCAG 3079
      :|||:|||||
Qy 960 IleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyIleAspTyrTyrValValArg 979
      |||||:|||||
Db 3080 ATTGCAGACTTCGCGCTTTCGCG-----GGAGAGAGGTTTATGTAGAGAG 3127
      |||||:|||||
Qy 980 GluProGlyGlnSerProIlePheTyrTyrAlaProGlySerLeuSerAspAsnIlePhe 999
      |||||:|||||
Db 3128 ACGATGGGGCTCTCCCTGCGCTGATGGCCATGTGATGCCATGACCTAGAGTCTAT 3187
      |||||:|||||
Qy 1000 SerArgIleSerAspValTyrPhePheGlyValValLeuTyrGluLeuPheThr----- 1017
      :|||:|||||
Db 3188 ACCACCAAGAGTATGTCTGTCTTGGAGTCTTCTTGGAGATAGTGAAGCTTGA 3247
      :|||:|||||
Qy 1018 -----TyrCysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetGly 1034
      :|||:|||||
Db 3248 GGTAACACCTACTGTGCGATGACCTGTGCC----- 3277
      |||||
Qy 1035 ProGluArgGluGlyProProLeuCysArgLeuLeuGluLeuAlaGluGlyArgArg 1054
      |||||
Db 3278 -----GAGCTCTATGATAAAGCTGCGCCAGGGCTTACCG 3310
      :|||:|||||
Qy 1055 LeuProProProProThrcysProThrcysGluValGlnGluLeuMetGlnLeuCysTyrPala 1074
      :|||:|||||
Db 3311 ATGAGAGAGCTCGAAGCTGAGCATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 3370
      :|||:|||||
Qy 1075 ProGluProHisAspArgProAlaPheAlaThrLeuSerProGlnLeuAspProLeu-Tr 1094
      |||||:|||||
Db 3371 GACGCTCCATGAGCAGCAGCCCTTGTGCCAGATGCGGTGACGATGAGCGCGCATGCTG 3430
      |||||:|||||
Qy 1094 PArg---GlyArgPro 1098
      |||||
Db 3431 GAAGCCAGGAAGGCT 3446
      |||||

RESULT 9
US-09-811-123-8
; Sequence 8, Application US/09811123
; Patent No. US20020001587A1
; GENERAL INFORMATION:
; APPLICANT: Sharon Erickson
; APPLICANT: Ralph Schwall
; APPLICANT: Mark Sliwowski
; TITLE OF INVENTION: METHODS OF TREATMENT USING ANTI-ERBB
; FILE REFERENCE: GENE 073A2
; CURRENT APPLICATION NUMBER: US/09/811,123
; CURRENT FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/238,327
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 09/602,530
; PRIOR FILING DATE: 2000-06-23
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 3768
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-811-123-8

Alignment Scores:
Pred. No.: 9,09e-40 Length: 3768
Score: 475.00 Matches: 261
Percent Similarity: 36.85% Conservative: 116
Best Local Similarity: 25.51% Mismatches: 337

```

```

Query Match: 8.11% Indels: 312
DB: 10 Gaps: 46
US-09-397-967-16 (1-1099) x US-09-811-123-8 (1-3768)
Qy 223 ProCysGlyArg---LeuProGlyArgProTyr---AlaIleMetAlaLysTyrIleLeu 240
      ||||| ||| ||||| |||||
Db 381 CCGTGTACAGAGGGGCTCCCGAGAGGCTCGGGAGGTGAGCTTGAACCCACAGCA 440
      :|||:|||||
Qy 241 AspLeuGluArg-----LeuHisProAlaIleThrGlnThrPheArgValGlyLeu 258
      |||||:|||||
Db 441 GATCTTGAAAGAGAGGGGCTTGAATCCAGCG----- 470
      |||||
Qy 259 ProGlyAlaGlnGluGluProGlyLeuLeuArgValAlaGlyAspAsnGlyIleProTyr 278
      |||||:|||||
Db 471 -----GAACCCCGACGCTGTGTA-----CCAGGA 494
      |||||
Qy 279 SerSerAsnAspGluLeuPheGlnThrPheCysAspPheProGluIleValAspValSer 298
      |||||:|||||
Db 495 CAC-----GATTTTGTGSA----- 509
      |||||
Qy 299 IleAsnGlnAlaProArgValGlyProAlaGlyIleHisArgLeuValThrValThrArg 318
      |||||:|||||
Db 510 --GGACATCTTCCACAAGAACACACAGCTGGCTGTCTAC----- 545
      |||||
Qy 319 MetAspGlyHis----- 322
      |||||
Db 546 ACTATATAGACACCAAGCGCTGCGGCTGCGACCCCTGTTCTCCGATGTGAAGGCTC 605
      |||||
Qy 323 ---IleLeuGlnAlaGluPheProGlyLeuProGluAlaLeuSerPheValAlaLeuVal 341
      :|||:|||||
Db 606 CCGCTGCTGGGAGAGAGATGCTGTGATGTCAGAGTC-CTGAGC----- 648
      :|||:|||||
Qy 342 AspGlyTyrPheArgLeuIleCysAspSerArgHisTyrPheCysGlyGluValAlaPro 361
      |||||:|||||
Db 649 -----CGACTGTGTGTGCGGTGCTGTGCTGCGCTGCAAGGGCCACTGCTCC 696
      |||||
Qy 362 ProArgLeuLeuGluGluGlu----- 368
      |||||
Db 697 ACTGACTGTGCTGATGAGCATGTGCTGCGGCTGCACAGGGCCCAAGCACTGTGACTGC 756
      |||||
Qy 369 -----AlaAspValCys-----HisGlyProIleThrLeu 378
      :|||:|||||
Db 757 CTGGCTGCTCCACTTCAACACACAGTGGCATCTGTGACTGCACTCCGACGCTGTGC 816
      :|||:|||||
Qy 379 AspPheAlaIleHisLysLeuLysAlaAlaGlySerLeuProGlyTyrTyrIleLeuArg 398
      :|||:|||||
Db 817 ACCTACACACACAGCACAGCTTTGAGTCCATGCCCATCCCGAGGGCCGCTAT----- 867
      :|||:|||||
Qy 399 ArgSerProGlnAspTyrAspSerPheLeuLeuThrAlaCysValGlnThrProLeuGly 418
      :|||:|||||
Db 868 -----ACATTCGGGGCAGCTGTGTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 915
      :|||:|||||
Qy 419 ProAspTyrLysGlyCysLeuIleArgGlnAspProSerGlyAlaPheSerLeuValGly 438
      :|||:|||||
Db 916 ACGGACGTGGGATCCCTCACCT--CGTCTGCCCT----- 950
      :|||:|||||
Qy 439 LeuSerGlnProHisArgSerLeuArgGluLeuLeuAlaLacysTyrPasn--SerGlyL 458
      :|||:|||||
Db 951 ----GCACACCAAGAGGTGACAGCAGGA-----TGAACACAGCCGCTG 992
      :|||:|||||
Qy 458 euArg-----ValAspGlyAlaAlaLeuTyrIleThrSerCysCys- 471
      :|||:|||||
Db 993 TGAGAAAGTCAGCAAGAGCCCTGTGCGCAGTGTGTATGTCTGGGCACTGGAGCACTTGC 1052
      :|||:|||||
Qy 472 -----AlaProArgProLysGlySerAsnLeuIleValValArgArg- 486
      :|||:|||||
Db 1053 AGAGGTGAGGAGGTATCCAGTGCATATATCCAGGACT--TTGCTGTGCTGAAGAGAT 1109
      :|||:|||||
Qy 487 -----GlyCysAsnProAlaProAlaP 494
      :|||:|||||
Db 1110 CTTTGGAGCCTTGCAATTCTGCCGAGAGACTTTGATGAGGAGACCACTTCACACTGC 1169
      :|||:|||||

```

QY 494 roglYcysSerProSer-----CysCysAlaLeu-----ThrGlnLeuSerP 508
 Db 1170 CCGGCTCCAGCCAGAGAGAGCTCCAGAGTGTAGAGACTGTGAGAGATCATCAGGTTAACT 1229
 QY 508 he-His-----ThrIlePro----- 512
 Db 1230 ATCAATCTCAGCATGGCCGAGACAGCTGCTGACCTCAGCTCAGCTCTCCAGAACCTGCAAGT 1289
 QY 513 -----ThrAspSerLeuGluTrp-----His 519
 Db 1290 AATCCGGGAGCAATCTGCACAAATGGCGGCTACTGCTGACCTGACCTGCAAGGCTGGGCAAT 1349
 QY 520 GlnAsn-LeuGlnHisGlySerPheThrIlePheArgGlySerArgArgGlyAlaVal 539
 Db 1350 CACCTGGCTGGGCTGGGCTGGCTGAGGAGAACTG-----GCCAGTGAGTGGCCCTCAT 1403
 QY 539 LAspGlyLutHisAspSerGluValLeuLeuValMetAspSerArgHisArgas 559
 Db 1404 CCACCAATACACCCACCTGCTTGGTGCACACGGTGGCCCTGGGACAGCTTTTGGGAA 1463
 QY 559 nCysMetGluSerPheLeuGlnAlaAlaSerLeuMetSerGlnValSerTyProHisIle 579
 Db 1464 C-----CCGACCA 1472
 QY 579 uValLeuLeuHisGlyValAlcysMetAlaGlyAspSerIleMetValGlnGluPheValTy 599
 Db 1473 AGCTCTGCTCCACACTGCCAACCCGCGCAGAGAC-----GAGTGTGTGGG 1517
 QY * 599 rLeuGlyAlaIleAspMetTyLeuArgGlyArgGlyHisLeuValSerAlaSerTrpTy 619
 Db 1518 CGAGGGCTGGGCTGGCACCACAGCTGTGGCCCGAGGGAC-----TGGTGGGG 1565
 QY 619 sLeuGlnValThrLys-----GlnLeuAlaTyAlaLeuAlaSerTyLeuGln 634
 Db 1566 TCCAGGGCCACCCAGCTGTGTCACTGACAGCAGCTTCCTGGGGCCAGAGTGGCTGA 1625
 QY 634 uAsp-----LysGlyLeuProHisGlyAsnValSerAlaArgLysValLeuLe 650
 Db 1626 GGAATCCCAAGTACTGACGGGCTCCCGAGGAGATGTGAATGCCAGGACACTGTTGGC 1685
 QY 650 uAlaArg-----GlnGlyLysAsp----- 656
 Db 1686 GTGCCACCTGAGTGCAGCCCCAGAAATGGCTCAGTGCAGCTGTTTGGACCGGAGCTGA 1745
 QY 657 -----GlyAsnProPheIleLysLeuSerAspPro-- 667
 Db 1746 CCAGTCTGTGGCTGTGCCACTATTAAGAGACCTCCCTTCTGCGTGGCCCGCCAG 1805
 QY 668 -GlyValSerProThrValLeuSerLeuLeuMetLeuThrAspArgIleProTrpValAl 687
 Db 1806 CGGTGTGAACCTGACCTCTCTCTACATGCCCATC-----TGGAACTT 1847
 QY 687 aProGlu-----CysLeuGlnGlnAlaGlnThrLeuCysLeuGln 700
 Db 1848 TCCAGATGAGGAGGGGCGCATGGCAGCCTTGGCCCATCACTCAGCCACTCTGTGTGA 1907
 QY 700 uAlaAspLysTrpGlyPheGlyAlaThrThrTrpGluValPheGlnArgGlyProAlaHis 720
 Db 1908 CCTGTGTGACAAGGCTGCCCGCC-----GAGCAGAGAGCCAGGCC 1949
 QY 720 sIleThrSerLeuGluProAlaLysLysLeuLysPheTyArgLysGlnGlyLeuP 740
 Db 1950 TCTGACGTCCATCTGCTCTGGCGTG-----GTTGGCATTTCTGCT 1988
 QY 740 cAlaLeuLysTrpThrGlnLeuAlaGlyLeuIleThrGlnCysMetAlaTyAspProGln 760
 Db 1989 GGTCTGTGCTTGGGGGTGCTTGGGATCTCATCAAG---CCACGGCAGCAAGAAT 2045
 QY 760 yArgArgProSerPheAlaGlnAlaIleLeuArgAspLeuAsnGlyLeuIleThrSerAspTy 780
 Db 2046 CCGGAAGTACACGATGCGGAGACTCTGCAAGAAACG----- 2082
 QY 780 rGluLeuLeuSerAspProThrProGlyIleProSerProArgAspGluLeuGlyValAl 800

Db 2083 -GAGCTGTGGAGACCCCTGACACTAGCCGAGGATGCCAACCA- 2127
 QY 800 aGlyAlaGlnLeuTyAlaCysGlnAspProAlaIlePheGlnGluArgHisLeuLysTy 820
 Db 2128 ---GGCAGATG-----CGATCCCTGAAGACAGGAGGTAGGAA 2165
 QY 820 rIleSerLeuGlnGlyLysGlyAsnPheLysSerValGlnLeuLysArgTyArgProLe 840
 Db 2166 GGTGAAGTGTGGATGTGGGCTTTTGGCAGACAGTCTCAAGGSCATGTGGATGCTCTGA 2225
 QY 840 uGlyAspAsnThrGlyProLeuValAlaValLysGlnLeu-----GlnHisSerValProAs 859
 Db 2226 TGGGAGAAATGTGAATTCGACATGCTGACAGCTTCAAGAGTGTGAGGAAACATCCSSCA 2285
 QY 859 rGlnGlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuHisSerAspPheI 879
 Db 2286 AGCCAAACAAAGAACTGTAGACGAAGCATAGATGAGCTGGGTGGGCTSSCATATGT 2345
 QY 879 eValLysTyArgGlyValSerTyGlyProGluArgGlnSerLeuArgLeuValMetI 899
 Db 2346 CTCCGCTTCTGGGCAATCTGCTG-----ACATCCAGGTCGACAGCTGTGACACA 2396
 QY 899 uTyLeuProSerGlyCysLeuArgAspLeuLeuGlnArgHisArgGly---LeuHisTh 918
 Db 2397 GCTTATGCCCTATGGCTGCTGTAGACATGTCCGGAACCAACCGGACAGCTGGGCTC 2456
 QY 918 rAspArgLeuLeuPheAlaTrpGlnIleCysLysGlyMetGlyTyTrpLeuGlyAlaAr 938
 Db 2457 CCAGGACCTGTGAACCTGTATGCGAGATGTCCAAAGGAGTACAGTACAGAGATGT 2516
 QY 938 gArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGlnSerGlnAlaHisVa 958
 Db 2517 GCGCTCTGTACACAGGAGCTTGGCGCTGGAGACGCTGGTCAAGAGTCCCAACCATGT 2576
 QY 958 lLysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlnGlyLysAspTyTrpValVa 978
 Db 2577 CAAAATTACAGACTTGGGCTGGCTGCTGGACATTCAGAGACAGACTACCATGC 2636
 QY 978 lArgGlyProGlnSerProIlePheTrpTyArgLarProGlnSerLeuSerAspAsnI 998
 Db 2637 A---GATGGGGCAAGTGGCCCATCAAGTGAATGGCGGTGATGCATTTCCCGCGGCG 2693
 QY 998 ePheSerArgGlnSerAspValTrpSerPheGlyValValLeuTyArgLysLeuPheThrTy 1018
 Db 2694 GTTCAACCAACAGATGATGTGTGACATGATGATGTGTGACGTGTGGAGCTGATACATTT 2753
 QY 1018 rCysAspLysSerCysSerProSerAlaGlnPheLeuArgMetGlyProGluArgGln 1038
 Db 2754 TGGGCGCAAA-----CCTTACGA 2771
 QY 1038 uGlyProProLeuCysArgLeuLeuGlnLeuAlaGlnGlyAlaGlyLeuProProP 1058
 Db 2772 TGGGATCCCAAGCCCGGAGATCTCTGACCTGTGTGAAAAAGGGGAGCGGCTGCCACCC 2831
 QY 1058 oProThrCysProThrGlnValGlnGlnLeuMetGlnLeuCysTrpAlaProGlnProH 1078
 Db 2832 CCCCATCTGCACCAATGATGTCTACATGATGATGATGATGATGATGATGATGATGATGATGAT 2891
 QY 1078 sAspArgProAlaPheAlaThrLeuSerProGlnLeuAspProLeuTrpArgGlyArgP 1098
 Db 2892 ATGTGCGCCAAAGATTCGCGGAGATGTGTGTGAAT--TCTCCCGCATGCCAGGAGACCC 2949
 QY 1098 o 1098
 Db 2950 C 2950
 RESULT 10
 us-09-811-115-2
 ; Sequence 2, Application us/09811115
 ; Patent No. US20020035736A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Erickson, Sharon


```

Qy 687 aProGlu-----CysLeuGlnGluAlaGlnThrLeuCysLeuG1 700
Db 1848 TCCAGATCAGAGCGCGCATCCACCCCTTGCCTCCATCACTGCACCCACCTCTGTGTGA 1907
Qy 700 uAlaAspLysTrpGlyPheGlyAlaThrThrTrpGluValPheGlnArgGlyProAlaH1 720
Db 1908 CCTGGATGACACAGGCTGCCCCGCC-----GAGCAGAGAGCCAGGCC 1949
Qy 720 sIleThrSerLeuLupProAlaLysLysLeuLysPheThrGlnLupSrgInGlyInLeuPr 740
Db 1950 TCTGACGTCCATCGTCTGGGGTG-----GTTGGCATCTCTCT 1988
Qy 740 cAlaLeuLysTrpThrGluLeuAlaGlyLeuIleThrGlnCysMetAlaTrpAspProG1 760
Db 1989 GGTGCTGTCTTGGGGGTGCTTGTGGATCTTCATCAAG---CGACGGCAGCAGACAGAT 2045
Qy 760 yArgArpProSerPhePheAlaIleLeuArgAspLeuAsnGlyLeuIleThrSerAspTy 780
Db 2046 CCGGAAGTACACGATCGGAGACCTGCTCAGGAAACG----- 2082
Qy 780 rGluLeuLeuSerAspProThrProGlyIleProSerProArgAspGluLeuCysValAl 800
Db 2083 -GAGCTGTGTGAGCGCGTGCACACCTAGGAGCGATGCCCAACG----- 2127
Qy 800 aGlyAlaGlnLeuTyraLysCysGlnAspProAlaIlePheGlnGluArgHisLeuLysTy 820
Db 2128 ---GCCGAGATG-----CGCATCCTGAAGAGCAGGAGCTAGGAA 2165
Qy 820 rIleSerLeuLeuGlyLysGlyAsnPheGlySerValGluLeuCysArgTyraAspProLe 840
Db 2166 GGTGAAGGTCTGTGATCTGGCCCTTTTGGCACGTCTACAAAGGCATCTGATCTCTGA 2225
Qy 840 uGlyAspAsnThrGlyProLeuValAlaValLysGlnLeu---GlnHisSerValProAs 859
Db 2226 TGGGAGAGATCTGAATAATTCAGTGGCCATCAAGTGTGAGGGGAAACATCCCCCA 2285
Qy 859 pGlnGlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuHisSerAspPheI1 879
Db 2286 AGCCACACAAAGAAATCTTAGACGAAGCATACGTGCTGCGCTCCCTCCATATGT 2345
Qy 879 eValLysTyraArgLysValSerTyraGlyProGlyArgInSerLeuArgValMetG1 899
Db 2346 CTCCCGCTTGTGGCATCTGCTG-----ACATCCAGGAGTGGCTGGTACACA 2396
Qy 899 uTyLeuProSerGlyCysLeuArgAspLeuLeuGlnArgHisArgGly---LeuHisTrp 918
Db 2397 GCTATATGCCCTATGAGCTGCTCTTAGACCAATGTCCGGAAAAACCGGAGCGCTGGCTC 2456
Qy 918 rAspArgLeuLeuLeuPheAlaIleTrpGlnIleCysLysGlyMetGluTyraLeuGlyAlaAr 938
Db 2457 CCAGGACCTGCTGAACCTGGTGTATGCGCAAGGGGATGAGCTACCTGGAGATGT 2516
Qy 938 gArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGlnSerGluAlaHisVa 958
Db 2517 GCGGCTGTACACAGGAGCTTGGCGCTCGAGACCTGCTGTCAAGATGCCAACCAATGT 2576
Qy 958 lLysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAspTyraValVa 978
Db 2577 CAATAATTCAGACTTGGCGCTGCGCTCGCTGACATTCGACGACAGATACCAGTGC 2636
Qy 978 lArgGluProGlyInSerProIlePheTrpTyraIleArgProGlyLysSerAspAsnI1 998
Db 2637 A---GATCGGGGCAAGTGGCCATCAAGTGAATGCGCTGAGTCCATCTCCGCGCGC 2693
Qy 998 ePheSerArgGlnSerAspValTrpSerPheGlyValValLeuTyraGluLeuPheThrTy 1018
Db 2694 GTTACACCAACAGATGATGTGTGTGAGTATGTGTGCTGTGTGGAGCTGATGACTTT 2753
Qy 1018 rCysAspLysSerCysSerProSerAlaGluPheLeuArgMetGlyProGlyArgG1 1038
Db 2754 TGGGGCCAAA-----CCTTAGCA 2771
Qy 1038 uGlyProProLeuCysArgLeuLeuGluLeuLeuAlaGluGlyArgArgLeuProPro 1058

```

```

Db 2772 TGGGATCCACCGCGGAGATCCCTGACCTGCTGGAAGAAAGGGGAGCGGCTGCCAGCC 2831
Qy 1058 pProThrCysProThrGlnValGlnGluMetGlnLeuCysTrpAlaProGluProH1 1078
Db 2832 CCCCATCTGCACCATTAATGATGTCTACATGATCATGTGTCAAATGTTGGATGATTGACTTGA 2891
Qy 1078 sAspArgProAlaPheAlaIleThrLeuSerProGlnLeuAspProLeuTrpArgGlyArgPr 1098
Db 2892 ATGTGCGCAAGATTCGCGGAGATTGTGTCTGAAT--TCTCCCGCATGGCCAGGAGCCCC 2949
Qy 1098 o 1098
Db 2950 c 2950

RESULT 11
US-09-877-177-11
; Sequence 11, Application US/09877177
; Publication No. US20020192652A1
; GENERAL INFORMATION:
; APPLICANT: Peter V. Danenberg et al.
; TITLE OF INVENTION: Method of determining Epidermal Growth
; FACTOR RECEPTOR: Factor Receptor and HER2-New Gene Expression
; FILE REFERENCE: 11220/120
; CURRENT FILING DATE: 2001-06-11
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 4530
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-877-177-11

Alignment Scores:
Pred. No.: 1,2e-39 Length: 4530
Score: 475.00 Matches: 261
Percent Similarity: 36.85% Conservative: 116
Best Local Similarity: 25.51% Mismatches: 337
Query Match: 8.11% Indels: 312
DB: 9 Gaps: 46

US-09-397-967-16 (1-1099) x US-09-877-177-11 (1-4530)
Qy 223 ProCysGlyArg---LeuProGlyArgProTyra---AlaLeuMetAlaLysTyraLeu 240
Db 531 CCCTGTACAGAGGCGCTCCCGACGAGCGCTGGGAGGTGCAGCTTCAAGCCTCAGACA 590
Qy 241 AspLeuGluArg-----LeuHisProAlaAlaIleThrGluThrPheArgValGlyLeu 258
Db 591 GATCTTGAAGAGGAGGGGTCTTGTATCCAGCG----- 620
Qy 259 ProGlyAlaGlnGluProGlyLeuArgValAlaGlyAspAsnGlyLeuProTrp 278
Db 621 -----GAAACCCCAAGCTCTGCTA-----CCAGGA 644
Qy 279 SerSerAsnAspGluLeuPheGlnThrPheCysAspPheProGluIleValAspValSer 298
Db 645 CAC-----GATTTTGTGGAA----- 659
Qy 299 lLeaSnGlnAlaProArgValGlyProAlaGlyLysHisArgLeuValThrValThrArg 318
Db 660 ---GGACATCTTCCACAGAACACACACAGCTGCTCTCAC----- 695
Qy 319 MetAspGlyHis----- 322
Db 696 ACTGATACACACCAACCGCTCTCGGCGCTGCCACCCCTGTTCTCCGATGTGAAGGCTC 755
Qy 323 ---lLeuGlnAlaGluPheProGlyLeuProGluAlaLeuSerPheValAlaLeuVal 341
Db 756 CCGCTGTGTGGGAGAGAGATTTCTGAGGATTTGTCAAGAC--CTGAGC----- 798
Qy 342 AspGlyTyraPheArgLeuIleCysAspSerArgHisTyraPheCysLysGluValAlaPro 361

```

Db 799 -----GCACACTGCTGTGGCGGTGGCTGTGCTGCCCGCTCAAGGGGCGACTGCC 846
OY 362 ProArgLeuLeuGluGluGlu-----
Db 847 ACTGACTGCTGCATGAGCAGGTGTGCTGCCGCTGCACGGGCCCCAAGCACTGTGACTGC 906
OY 369 -----AlaAspValCys-----HisGlyProIleThrLeu 378
Db 907 CTGGCGCTGCCTCACTTCACACACAGTGGCATGTGTGAGCTGCATGCCCTCCAGCCCTGTGC 966
OY 379 AspPheAlaIleHisLysLeuLysAlaAlaGlySerLeuProGlyThrTyrlleLeuArg 398
Db 967 ACCTACACACAGACACAGCTTTGAGTCCATCCCAATCCAGGGCCGGTAT----- 1017
OY 399 ArgSerProGlnAspTyrAspSerPheLeuLeuThrAlaSerValGlnThrProLeuGly 418
Db 1018 -----ACATTCGGCGCGCGCTGTGTGACTGCTGTGCTTCACAACTACTTTCT 1065
OY 419 ProAspTyrLysGlySerLeuIleArgGlnAspSerGlyAlaPheSerLeuValGly 438
Db 1066 ACGGACGTGGATTCCTGCACCT-CTGTGCGCCCT----- 1100
OY 439 LeuSerGlnProHisArgSerLeuArgGluLeuLeuAlaLacystTrpAsn--SerGlyL 458
Db 1101 ---GCACACACCAAGAGGTGACAGCAGAGA-----TGCAACACAGCGGTG 1142
OY 458 euArg-----ValAspGlyAlaAlaLeuTyrlleThrSerCysCys- 471
Db 1143 TGAGAGGTGACAGAGCGCTGTGCGCGAGTGTGTATGTGTGGCGCATGGAGCACTTGGCG 1202
OY 472 -----AlaProArgProLysGlySerAsnLeuIleValIleArgArg- 486
Db 1203 AGAGGTGAGGCGAGTTACCAAGTGCATATCCAGAGCT--TTGCTGCTGCACAGAGAT 1259
OY 487 -----GlyCysAspProAlaProAla 494
Db 1260 CTTTGGGAGCGCTGCATTTCTGCCGAGAGCTTTGATGGGAGCCAGCTTCACAACTGC 1319
OY 494 roGlyCysSerProSer-----CysCysAlaLeu-----ThrGlnLeuSerP 508
Db 1320 CCGGCTCCACCGACAGAGCTCCACAGTGTGTGAGACTGTGAGAGAGATACAGAGTTACT 1379
OY 508 he-His-----ThrIlePro----- 512
Db 1380 ATACATCTCAGCATGGCCGAGACAGCTGCCAGCTCAGCGCTTCCAGAACCTGCAAGT 1439
OY 513 -----ThraspSerLeuGluTrp-----His 519
Db 1440 AATCCGGGGGAGCAATTTTGCACAAATGGCGCTACTGCTGACCTGCAGAGCGCTGGCAT 1499
OY 520 GlnAsn-LeuGlyHisGlySerPheThrLysIlePheArgGlySerArgArgGluVal 539
Db 1500 CAGCTGGCTGGGGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1553
OY 539 LaspGlyLysThrHisAspSerGluValLeuLeuLysValMetLaspSerArgHisArg 559
Db 1554 CCACCATTAACACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1613
OY 559 ncysMetGluSerPheLeuGluAlaAlaSerLeuMetSerGlnValSerTyrProHis 579
Db 1614 C-----CGCACCA 1622
OY 579 uValLeuLeuHisGlyValCysMetAlaGlyAspSerIleMetValGlnGluPheVal 599
Db 1623 AGCTGCTGCTCCACACATGCCAAGCGGCGCAGAGAC-----GAGTGTGGGG 1667
OY 599 rLeuGlyAlaIleAspMetTyrLeuArgLysArgGlyHisLeuValSerAlaSerTrp 619
Db 1668 CGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1715
OY 619 sLeuGlnValThrLys-----GlnLeuAlaIleTyrAlaLeuAsnTyrTrpLeuG 634
Db 1716 TCCAGGGGCCACCCAGTGTGTCAATGCAGCCAGCTTCTTGGGGCCAGAGTGTGTGA 1775
OY 634 uAsp-----LysGlyLeuProHisGlyAsnValSerAlaIleArgLysValLeu 650
Db 1776 GGAATGCCGAGTACTGAGGGGCTCCCGAGGAGTATGTGATGCCAGGACTGTGGC 1835
OY 650 uAlaArg-----GluGlyGlyAsp----- 656
Db 1836 GTGCCACCTGATGTCAGCCCGCAGAAATGGCTGATGACCTGTTTGGACCGGAGCTGA 1895
OY 657 -----GlyAsnProPheIleLysLeuSerAspPro-- 667
Db 1896 CCATGTGTGGCCCTGTGCCCACTAATTAAGACCTCCCTTGTGCGCTGGCCCGCTCCAG 1955
OY 668 -GlyValSerProThrValLeuSerLeuGlnMetLeuThrAspArgLleProThrValAl 687
Db 1956 CGGTGTAAACCTGACCTCTCTCTACATGCCATC-----TGGAAGTT 1997
OY 687 aProGlu-----CysLeuGlnGluAlaGlnThrLeuLeuG 700
Db 1998 TCCAGATGAGGAGGCGCCATGCCAGCTTGGCCCATGACAGTGCACCACTCTGTGGA 2057
OY 700 uAlaAspLysTrpLysPheGlyAlaThrThrProGluValPheGlnArgGlyProAla 720
Db 2058 CTTGATGACAGAGGCTGCGCCGCC-----GAGCAGAGAGCCAGCC 2099
OY 720 sIleThrSerLeuLeuProAlaLysLysLeuLysPheTyrlleLysAspGlnGlyLeuPr 740
Db 2100 TCTGACCTGCATCTCTGTGGGTG-----GTGGCATTTGCT 2138
OY 740 oAlaLeuLysTrpThrLeuLeuAlaGlyLeuIleThrGlnCysMetAlaIleAspProG 760
Db 2139 GGTGCTGTCTTGGGGGTGTGCTTGGGATCTCATCAG--CGACGCGACAGAAAGAT 2195
OY 760 YArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIleThrSerAspTy 780
Db 2196 CCGGAGTACACAGATGGGAGACTGCTGCAGAAAG-- 2232
OY 780 rGluLeuLeuSerAspProThrProGlyIleProSerProArgAspGluLeuCysValAl 800
Db 2233 -GAGCTGTGTGAGCGCTGCAGACCTGACGAGCGATCCCAACAG----- 2277
OY 800 aglyAlaGlnLeuTyrlleCysGlnAspProAlaIlePheGlnGluArgHisLeuLys 820
Db 2278 ---GCCCAGATG-----CGATCTCGAAGAGAGCGAGCGAGAGAA 2315
OY 820 rIleSerLeuLeuGlyLysGlyAsnPheGlySerValGluLeuCysArgTyrAspProLe 840
Db 2316 GGTAAAGTGTGATGTGATGTGCTTGGCAGCTGCTACAGGCGCATGTGATCCCTGA 2375
OY 840 uGlyAspAsnThrGlyProLeuValAlaValLysGlnLeu--GlnHisSerValProAs 859
Db 2376 TGGGGAATATGTGAATAATTCAGAGCATCAAAATGTTGAGGAAACACATCCCAA 2435
OY 859 pGlnGlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuHisSerAspPhe 879
Db 2436 AGCCAAACAAAGAAATCTTAAGCAGAGCATACGTGATGCTGTGGCTGCCCATATGT 2495
OY 879 eValLysTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuArgLeuValMet 899
Db 2496 CTCCCGCTTCTGGCATCTGCTG-----ACATTCACGGTGCAGCTGTGACACA 2546
OY 899 uTyrlleuProSerGlyCysLeuArgAspLeuGlnArgHisArgGly--LeuHisThr 918
Db 2547 GCTTATGCCCTATGCTGCTCTTAAGCAATGTCCGGAAACCGGAGCGCTGGGCTC 2606
OY 918 rAspArgLeuLeuLeuPheAlaTrpGlnIleCysLysGlyMetGluTyrlleuGlyAla 938
Db 2607 CCAGGACCTCTGTAACAGGTATATGATGCAAGGGATGAGCACTCCGAGGAGATGT 2666
OY 938 gArgCysValHisArgAspLeuAlaIleArgAsnIleLeuValGluSerGluAlaHisVa 958
Db 2667 GCGGCTGTATACAGGAGACTTGGCGCTCGAAGAGTGTGTCAAGAGTCCCAACACATGT 2726

QY 358 uValAlaProArArgLeuGluGlu----- 368
 |||
 Db 2417 GCCACTGCCACGTACGTGCGCATGAGACAGTGTGCGCGGTGCACGGGCCCAAGCA 2476
 QY 369 -----AlaSpValCys-----HisGlyPyr 375
 Db 2477 CTCTGACTGCGTGGCTGCTCCACTTCACACACAGTGGATCTGTGAGCTGCACTGCC 2536
 QY 375 oIleThLeuAspPheAlaIleHisLysLeuValAlaGlySerLeuProGlyThrTy 395
 |||
 Db 2537 AGCCCTGGTCACTACACACACAGACGTTTGAGTCCATGCCCAATCCGAGGGCCGTA 2596
 QY 395 rIleLeuArgArgSerProGlnAspTyrAspSerPheLeuLeuThrAlaCysValGlnTh 415
 |||
 Db 2597 T-----ACATTGGCGCGAGCTGTGTGCTGCTGCTGCTCCCTCA 2635
 QY 415 rProLeuGlyProAspTyrLysGlyCysLeuIleArgGlnAspSerGlyAlaPheSe 435
 |||
 Db 2636 CTACCTTTCTACGACGTGGAGTCTGCACCT--CGTCTGCCCT----- 2680
 QY 435 rLeuValGlyLeuSerGlnProHisArgSerLeuArgGluLeuLeuAlaIaCysTrpAs 455
 |||
 Db 2681 -----GCACACCAAGAGTGACAGCAGAGA-----TGGA 2712
 QY 455 n--SerGlyLeuArg-----ValAspGlyAlaAlaLeuTyrLeuThr 468
 |||
 Db 2713 CACAGCGGTGAGAAAGTGCAGCAGACGCTGTGCGCGAGTGTCTATGTGTGGGCATGG 2772
 QY 469 SerCysCys-----AlaProArgProLysGlyLysSerAsnLeuIleVal 483
 |||
 Db 2773 AGCACTTGGCAGAGGTGAGGGCACTTACAGTGCCTCAATTCAGGAGCT--TTCTGCT 2829
 QY 484 ValArgArg-----GlyCysAsnPro 490
 |||
 Db 2830 GCAAGAAGATCTTGGAGCGCTGCGCATTTCTGCCGAGAGCTTGTGATGGGACCCAGCT 2889
 QY 491 AlaProAlaProGlyCysSerProSer-----CysCysAlaLeu-----Thr 504
 |||
 Db 2890 CCAACACTGCCCGCTCCAGCCAGCAGAGCTCCAGAGTGTGAGACTCTGAGAGACATCA 2949
 QY 505 GlnLeuSerPhe--His-----ThrIlePro-- 512
 |||
 Db 2950 CAGGTACCTATACATCTACAGCATGCGCGGACACCTGCTGACCTCAGCGCTTCCAGA 3009
 QY 513 -----ThrAspSerLeuGluTrp----- 518
 |||
 Db 3010 ACCTGCAAGTAATCCGGGAGAGATTTCTGCACATGCGCGCTACTGCTGACCTGCAG 3069
 QY 519 -----HisGlnAsn--LeuGlyHisGlySerPheThrLysIlePheArgGlySerArg 536
 |||
 Db 3070 GGCTGGGCAATCAGCTGTGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3123
 QY 536 rGluValValAspGlyGlyThrHisAspSerGluValLeuLeuLysValMetAspSer 556
 |||
 Db 3124 TGGCCCTCATCCACCAATACACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3183
 QY 556 rGlnHisArgAsnCysMetGluSerPheLeuGlnAlaAlaSerLeuMetSerGlnValSer 576
 |||
 Db 3184 TCTTTCGAGC----- 3194
 QY 576 yrrProHisLeuValLeuLeuHisGlyValCysMetAlaGlyAspSerIleMetValGln 596
 |||
 Db 3195 --CCGACCAAGCTCTCTCCACACTGCCAACCGGCCAGAGAGAC-----G 3237
 QY 596 luePheValTyrLeuGlyAlaIleAspMetTyrLeuArgLysArgGlyHisLeuValSer 616
 |||
 Db 3238 AGTGTGTGGCGGAGGCGCTGCGCTGCGCACACAGCTGTGCGCGCGGAGGCGAC----- 3287
 QY 616 lAserTrpLysLeuGlnValThrLys-----GlnLeuAlaTyrAlaLeu 631
 |||
 Db 3288 --TGCTGGGTCCAGGGCCACACCACTGTGTCAACGTGACGACGACTGCTGCGGGCCAGG 3345
 QY 631 snTyrLeuGluAsp-----LysGlyLeuProHisGlyAsnValSerAlaArgL 647

Db 3346 AGTCGTGAGAGATGCGAGTACTGACAGGGGCTCCCGAGGAGATATGTATCATCAGGC 3405
 QY 647 yValLeuLeuAlaArg-----GlnGlyGlyAsp----- 656
 Db 3406 ACTGTTCGCGTGCACCTGAGTGTGACGCCCAAGTGTGTGACACTGTTTGGAC 3465
 QY 657 -----GlyAsnProPheIleLysLeus 665
 Db 3466 CGAGGCTGACCACTGTGTGCGCTGTGCCACTTAAAGAGACCTCCCTTGTGCGTGGCC 3525
 QY 665 eAspPro--GlyValSerProThrValLeuSerLeuGluMetLeuThrAspArgIleP 684
 |||
 Db 3526 GCTGCCCGGCGGTGTGAACCTGACCTCTCTCATAGCCCATC----- 3569
 QY 684 rOTrPValAlaProGlu-----CysLeuGlnGluAlaGlnThrL 697
 |||
 Db 3570 --TGGAAGTTTCCAGATGAGAGAGGGCCCATGCCACTTGCCTCATGACCCACT 3627
 QY 697 euCysLeuGlnAlaAspLysTrpGlyPheGlyAlaThrTrpGluValPheGlnArg 717
 |||
 Db 3628 CCTGTGTGAGCTGTGATGACAGGGGCTGCCCGC-----GAGCAGA 3669
 QY 717 lYProAlaHisIleThrSerLeuGluProAlaLysLysLeuLysPheTyrGluAspGln 737
 |||
 Db 3670 GAGCCAGCCCTGACCTGACCTGCTCTCTGCGGTG-----GTG 3708
 QY 737 lGlnLeuProAlaLeuLysTrpThrGlnLeuAlaGlyLeuIleThrGlnCysMetAla 757
 |||
 Db 3709 GCATTCTGCTGTGTGTGTGCTGCGGGGTGTGCTTGGGATCTTATCAAG--CGAGGC 3765
 QY 757 yAspProGlyArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIle 777
 |||
 Db 3766 AGCAGAAGATCCGAGATACAGATGCGAGATCTGACAGAAAG----- 3812
 QY 777 hrSerAspTyrGluLeuLeuSerAspProThrProGlyIleProSerProArgAspGlu 797
 |||
 Db 3813 -----GACCTGTGTGAGCGCGCTGACACTTACCGGAGCAGTCCCAACAG-- 3857
 QY 797 euCysValAlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGlnGluArg 817
 |||
 Db 3858 -----GCGCAGATG-----CGATCTCTGAAGAAGACG 3885
 QY 817 lIleuLysTyrIleSerLeuLeuGlyLysGlyAsnPheGlySerValGluLeuCysArg 837
 |||
 Db 3886 AGCTGAGGAAGATGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3945
 QY 837 yAspProLeuGlyAspAsnThrGlyProLeuValAlaValLysGlnLeu--GlnHis 856
 |||
 Db 3946 GGATCCCTGATGGGAGAAATGTGAATTCAGATGCCATCAAGTGTGAGGAAACA 4005
 QY 856 eValProAspGlnGlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuHis 876
 |||
 Db 4006 CATCCCCCAAGCAACAAAGAAATCTTAAGCAAGCATACGTATGCTGTGGCT 4065
 QY 876 eAspPheIleValLysTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuArg 896
 |||
 Db 4066 CCCCATATGCTCCCGCTTCTGCGCATCTGCCG-----ACATCCACGCTGCAGC 4116
 QY 896 euValMetGluTyrLeuProSerGlyCysLeuArgAspLeuGlnArgHisArgGly- 915
 |||
 Db 4117 TGTGTACACAGCTTATGCTCATGCTGCTCTTGAACATGTCTGGGAAACCGCGGAC 4176
 QY 916 --LeuHisAspArgLeuLeuLeuPheAlaTrpGlnIleCysLysGlyMetGluTyr 935
 |||
 Db 4177 GCTGTGGCTCCAGAGCTGCTGTGAACGTGTATGAGATGTGCCAAGGAGTATGAGTACC 4236
 QY 935 euGlyAlaArgArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGluSer 955
 |||
 Db 4237 TGGAGGATGTGCGGCTGTACACAGGACCTTGGCGGAAAGTGTGTGTAAGAGTC 4296
 QY 955 lAlaHisValLysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAsp 975

[illegible][illegible]

QY 995 erAspAnIlePheSerArgInSerAspValTPrSerPheGlyValValLeuTyrGlu 1015
 Db 4414 TCCSCCGGCGGATCCACACAGATGTGTGAGTATGACTGTGGAGC 4473
 QY 1015 euphethrTyrCysAspLysSerSerProSerAlaGluPheLeuArgMetGlyP 1035
 Db 4474 TGAGACTTTTGGGCCAAA----- 4493
 QY 1035 roGluAArgGluGlyProProLeuCysArgLeuLeuGluLeuAlaGluGlyArg 1055
 Db 4494 --CCTTACGATGGGATCCACCGCGGAGATCCCTGACCTCTGMAAAGGGAGCGC 4551
 QY 1055 eupProProProThCysProThGluValGlnGluLeuMetGlnLeuCysTrpAlaP 1075
 Db 4552 TGCCCGACGCCCATCTGTGCACCATTTGATGTCTACATGATCATGTCMAATGTTGATGA 4611
 QY 1075 roGluProHIsAspArgProAlaPheAlaThrLeuSerProGlnLeuAspProLeuTrpA 1095
 Db 4612 TTGACTGTGAATGTGGGACAGATTCCGGAGTTGCTGTGAAT--TCTCCCGCATGGC 4669
 QY 1095 rGgLyArGPro 1098
 Db 4670 CAGGACCCCC 4680

RESULT 14
 US-09-854-356-9
 ; Sequence 9, Application US/09854356
 ; Patent No. US20020177567A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cheever, Martin A.
 ; APPLICANT: Corixa Corporation
 ; APPLICANT: SmithKline Beecham Biologicals S. A.
 ; TITLE OF INVENTION: HER-2/neu Fusion Proteins
 ; FILE REFERENCE: 014058-009810PC
 ; CURRENT APPLICATION NUMBER: US/09/854,356
 ; CURRENT FILING DATE: 2001-05-09
 ; PRIOR APPLICATION NUMBER: US 09/493,480
 ; PRIOR FILING DATE: 2000-01-28
 ; PRIOR APPLICATION NUMBER: US 60/117,976
 ; PRIOR FILING DATE: 1999-01-29
 ; NUMBER OF SEQ ID NOS: 26
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 9
 ; LENGTH: 3768
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(3768)
 ; OTHER INFORMATION: human HER-2/neu protein
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)..(1959)
 ; OTHER INFORMATION: extracellular domain (ECD) of human HER-2/neu
 ; NAME/KEY: misc_feature
 ; LOCATION: (2026)..(3765)
 ; OTHER INFORMATION: intracellular domain (ICD) of human HER-2/neu
 ; NAME/KEY: misc_feature
 ; LOCATION: (2968)..(3765)
 ; OTHER INFORMATION: phosphorylation domain (PD) of human HER-2/neu
 ; NAME/KEY: misc_feature
 ; LOCATION: (2968)..(3144)
 ; OTHER INFORMATION: preferred portion of the phosphorylation domain
 ; OTHER INFORMATION: (delta PD) of human HER-2/neu
 US-09-854-356-9

Alignment Scores:
 Pred. No.: 1.16e-39 Length: 3768
 Score: 474.00 Matches: 261
 Percent Similarity: 36.85% Conservative: 116
 Best Local Similarity: 25.31% Mismatches: 337
 Query Match: 8.09% Indels: 312
 Db: 9 Gaps: 46

US-09-397-967-16 (1-1099) x US-09-854-356-9 (1-3768)
 QY 223 ProCysGlyArg---LeuProGlyArgProTyr---AlaLeuMetAlaLysTyrLeu 240
 Db 381 CCTGTCTACAGGGGCTCTCCAGGAGGCTCGGAGGCTGACCTTCACACCTTCACAGA 440
 QY 241 AsPLeuGluArg-----LeuHIsProAlaAlaThrTrpGluThrPheArgValGlyLeu 258
 Db 441 GATCTTGAAAGAGGGGCTGTGATCCAGC----- 470
 QY 259 ProGlyAlaGlnGluGlyProGlyLeuLeuArgValAlaGlyAspAsnGlyLeuProTrp 278
 Db 471 -----GAACCCCGACGCTCTGCTA-----CCAGGA 494
 QY 279 SerSerAsnAspGluLeuPheGlnThrPheCysAspPheProGluIleValAspValSer 298
 Db 495 CAC-----GATTTGTGGA----- 509
 QY 299 IleAsnGlnAlaProAlaArgValGlyProAlaGlyGlnHIsArgLeuValThrValThrArg 318
 Db 510 --GGACATCTTCCACAGAACACACACCTGGCTCTAC----- 545
 QY 319 MetAspGlyHIs----- 322
 Db 546 ACTGATAGACACACACACCGCTCTCGGCGCTGCACCCCTTTCTCGATGTGAAGGCTC 605
 QY 323 --IleLeuGlnAlaGluPheProGlyLeuProGlnAlaLeuSerPheValAlaLeuVal 341
 Db 606 CCGCTGCTGGGAGAGAGATTCTGAGAGATTGTACAGAC--CTGAC----- 648
 QY 342 AspGlyTyrPheArgLeuIleCysAspSerArgHIsTyrPheCysAlaGlyValAlaPro 361
 Db 649 -----CCGACTGCTGTGCCGCTGCTGCTGCTGCTGCTGCAAGAGGCGCACTGCC 696
 QY 362 ProArgLeuLeuGlnGlu----- 368
 Db 697 ACTGACTGCTGCCATGACAGCTGTGCTGCCGCTGCACGGGCCCAACACTGTGACTGC 756
 QY 369 -----AlaAspValCys-----HIsGlyProIleThrLeu 378
 Db 757 CTGGCTGCCCTCCACTTCACACACAGAGTGCATCTGTGAGTGCATGCCAGCCCTGGTTC 816
 QY 379 AspPheAlaIleHIsLysLeuLysAlaAlaGlySerLeuProGlyThrTyrIleLeuArg 398
 Db 817 ACCTACACACACACACACCTTTGAGTCCATGCCCAATCCCGAGGCGGCTAT----- 867
 QY 399 ArgSerProGlnAspTyrAspSerPheLeuLeuThrAlaCysValGlnThrProLeuGly 418
 Db 868 -----ACATTCGGCGCCAGCTGTGTGACTGCTGCTCCCTACACACTTACTTCT 915
 QY 419 ProAspTyrLysGlyCysLeuIleArgGlnAspProSerGlyAlaPheSerLeuValGly 438
 Db 916 ACGGACGTGGGATCTGTGACCTT--CGTCTGCCCT----- 950
 QY 439 LeuSerGlnProHIsArgSerLeuArgGluLeuLeuAlaAlaCysTrpAsn--SerGlyL 458
 Db 951 ---GCACAAACAGAGGTGACAGACAGGA-----TGAACACACAGCGGTG 992
 QY 458 euArg-----ValAspGlyAlaAlaLeuTyrLeuThrSerCysCys- 471
 Db 993 TGAGAAGTGACAGAACCCCTGTGCCGAGTGTGCTGATGTGTGGGCATGAGACACTTTCG 1052
 QY 472 -----AlaProArgProLysGluLysSerAsnLeuIleValAlaArgArg- 486
 Db 1053 AGAGGTGAGGCGAGTTACCACTGCAATATCCAGAGAT--TTGCTGCTGCACAGAAAT 1109
 QY 487 -----GlyCysAsnProAlaProAlaP 494
 Db 1110 CTTTGGAGGCTGGGATTTGTGCGGAGAGCTTTGATGAGGAGCCAGGCTCCACACATGCG 1169
 QY 494 roGlyCysSerProSer-----CysCysAlaLeu-----ThrGlnLeuSerP 508

Db	1170	CCGGCTCCACGACAGACAGCTCCAAAGTGTTTAGACTCTTGGAAGATCAACAGTTAACT	1229
OY	508	he-His-----ThrIlePro-----	512
Db	1230	ATACAAATCTGAGCATGCGCGGACAGCTGACTGACCTGACGCTGCTCCAGAACTGCAAGT	1289
OY	513	-----ThrasperLeuGIuTP-----His	519
Db	1290	AATCCGGGACGAAATTCGACACATAGCCCTTACTGCTGACCTCGAAGGCTGGGCAT	1349
OY	520	GIuAsn-LeuGIuHisGIySerPheThrIysIlePheArgGIySerArgArgIuVala	539
		:::	
Db	1350	CAGCTGGCTGGGGCTCGCTGCTACCTAGAGGAATG-----GGCAGTGCAGCTGGGCCCTCAT	1403
OY	539	IAspGIyGIuThHisAspSerGIuValLeuLeuIysValMetAspSerArgHisArgas	559
Db	1404	CCACCAATAACCCACCTCTGCTGTCACACAGCTGGCCCTGGACCAACCTCTTTCGGA	1463
OY	559	nCysMetGIuSerPheLeuGIuValaIaSerLeuMetSerGIuValSerIyPProHisIe	579
Db	1464	C-----CCGACCA	1472
OY	579	uValLeuLeuHisGIyValCysMetIlaGIyAspSerIleMetValGIuGIuPheValIy	599
Db	1473	AGCTGTGCTCCACACTGCCAACCGCGCAGAGAC-----GAGTGTGTGGG	1517
OY	599	rLeuGIyAlaIleAspMetIyLeuArgIyLysArgGIyHisLeuValSerAlaSerIleTPly	619
Db	1518	CGAGGCGCTGGCGCTGCCACAGCTGTGGCGCCGAGGGCAC-----TCTTGGG	1565
OY	619	sLeuGIuValThrIys-----GIuLeuAlaIyAlaLeuAsnIyLeuGI	634
Db	1566	TCCAGGCGCCACCCACCATGTGTCACTGACAGCCATTCTTCGGGGCAGAGATGGCTGA	1625
OY	634	uAsp-----LysGIyLeuProHisGIyAsnValSerAlaArgIyValLeuIe	650
Db	1626	GGAAATGCCAGTACTGCAAGGGCTGCCCGAGGAGTATGTAAATCCAGGACACTGTTTGCC	1685
OY	650	uAlaIArg-----GIuGIyGIyAsp-----	656
Db	1686	GGCCACCTGTAAGTTCACCCCCAGAAATGGCTCAGTACCTGTTTGGACCGAGGCTGA	1745
OY	657	-----GlyAsnProPheIleIyLeuSerAspPro--	667
Db	1746	CCAGTGTGTGGCTGTGCCCACTATAGGAACCTCTCTCTGCTGGCGCCGCTGCCCCAG	1805
OY	668	-GIyAlaSerProThrValIleuSerLeuGIuMetIleuThAspArgIleProTPyValAl	687
Db	1806	CGGTGTGAACCACTGACCTCTCTCCATAGGCCATC-----TGAAGT	1847
OY	687	arProIu-----CysLeuGIuValaGIuThrLeuGIuLeuGI	700
Db	1848	TCCAGATGAGGAGGCGCATGCCAGCCTTGGCCCCATCACTGCACCCACACTCCGTGTGA	1907
OY	700	uAlaAspIySTrPGIyPheGIyAlaThrThrTrpGIuValPheGIuArgGIyProAlaHI	720
Db	1908	CCTGATGACAAAGCGCTGCCCCGCC-----GAGCAGAGAACCAAGCC	1949
OY	720	sIleThrSerLeuGIuProAlaIaLysIyLeuIysPheIyGIuAspGIuGIyLeuIePr	740
Db	1950	TCTGACGTCATCATCTCTGGCGGTG-----GTTGGCATTCGCT	1988
OY	740	oAlaLeuIySTrPThGIuLeuAlaGIyLeuIleThrGIuGIyMetAlaIyTrsPProGI	760
Db	1989	GGTCGTGGCTTGGGGGTGGCTTGGATTCCTCATCAAG--CGACGCGACGAGAAGAT	2045
OY	760	YArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGIyLeuIleThrSerAspIy	780
Db	2046	CCGGAAGTACAGATTCGGGAGACTGCTGCAGGAAGG-----	2082
OY	780	rGIuLeuLeuSerAspProThrProGIyIleProSerProArgAspGIuLeuGIyValAl	800
Db	2083	-GAGCTGTGGAGCCGCTCAACACTGTAGCGAGAGATGCCAACACAG-----	2127

QY	800	acglaagcaglnleutryralacysglnspproalalierhpeglucgluaarghlsleuylsty	820
		:++	
Db	2128	---CGCGAGATG-----CGGATCTGAAAGAGCGAGCTTGAGGAA	2165
QY	820	rilerseuleuendilysglsyasnphegyservalglnleucysargtryaspProle	840
		:	
Db	2166	ggTGAAGGTGCTTGATCTGGCCCTTTTGGCAAGCTCTCAAGGGCATCTGGATCCCTGA	2225
QY	840	ugllyaspasnThrlygProleuValAlavalysglnleu---GlnHisserValProas	859
		:	
Db	2226	TGGGGAGATATGTAATTCAGTCCAGTGCACATCAAGGTGTGAGGAGAAACACATCCCCAA	2285
QY	859	pGLnlglnhArgphheglInmrGluIleGlnIleleuysalaleuHisserAspPheII	879
		:	
Db	2286	AGCCACAAGAAATAATCTTAACACAGACATACGATGAGCTGTGTGGGCTCCCATATGT	2345
QY	879	eValIlystrArggllyalysertgrygProgllyArgInserLeuArgleuValMetG1	899
		:	
Db	2346	CTCCGGCTTGTGGGCAATCTGGCTG-----ACATCCAGGTCGACGTGTGACACA	2396
QY	899	uTytleuProserGlyCysLeuArgaspLeuGlnArgHisargGly---LeuHisTh	918
		:	
Db	2397	GCTTATGCTCATATGGCTGCTCTTAAACCATGTCGGGAAAAACCGGACCCCTGGGCTC	2456
QY	918	rAspArgleuIleuIleuIleuIleArgpGlnIleCyslysglyMetG1uTytleuGlyAlaAr	938
		:	
Db	2457	CGAGGACTGTGTACTGTGGTGTGCAAGATGCGCAAGGGGATGAGACTACTGGAGAGATGT	2516
QY	938	gArgCysValHisArgaspLeuAlaIlaArgsnIleleuValGInserGlnAlaHisVa	958
		:	
Db	2517	GCGGCTGTGTACAGGGAGACTTGGCCGCTGGAAACGTGTCGTCAAGAGATCCCAACCATGT	2576
QY	958	lLysIleAlaAspPheglyleuAlaLysleuIleuProleuGlyLysAspTyrTyrrValVa	978
		:	
Db	2577	CAAAATTAACAAGACTTCGGCTGCTGCTGCTGCGACATTGACGAGACAGATCCATGTC	2636
QY	978	lArgGlnProgllyGlnserProIlePheTrpTyrrAlaProGlnIserLeuSerAspAsnI1	998
		:	
Db	2637	A---GATGGGGGCAAGGTGCCCTCAAGTGGATGGCCGCTGGAGTGCATTTCCCGGGCG	2693
QY	998	ePheSerArgGlnSerAspValTrpserPheglyValValLeuTyrgluIleuPheTrhTy	1018
		:	
Db	2694	GTTACCCACACAGAGATGTGTGGAGTTATGGTGTGACTGTGGAGAGCTGATGACTTT	2753
QY	1018	rCysAspLysserCysSerProserAlaGluPheleuArgMetMetGlyProGlnArgG1	1038
		:	
Db	2754	TGGGGCCAAA-----CCTTACGA	2771
QY	1038	uG1ProProleuCysArgleuIleuGlnleuAlaGlnGlyArgArgleuProProPr	1058
		:	
Db	2772	TGGGATCCACACCCGGAGATCCCTTACCTGCTGGAAAAAGGGAGGGCGGCCCCAGCC	2831
QY	1058	oProThrCysPProThrgluValGlnGluIleuMetGlnleuCystrAlaIaProGluProH1	1078
		:	
Db	2832	CCCCATCTGGCACCATTTGATCTTACATGTACATGTCATAGTTCATGATGTGACTGTGA	2891
QY	1078	sAspArgProAlaPheAlaTrhIleuSerProGlnIleuAspProleuTrpArgGlyArgTr	1098
		:	
Db	2892	ATGTGGCCCAAGATTCGGGAGTGTGTCTGAAT--TCCTCCGATGGCCAGGAGACCC	2949
QY	1098	o 1098	
Db	2950	C 2950	

RESULT 15
US-09-930-125-1
; Sequence 1, Application US/09930125
; Publication No. US20020193329A1
; GENERAL INFORMATION:
; APPLICANT: Hand-Zimmerman, Susan
; APPLICANT: Cheever, Martin A.
; APPLICANT: Foy, Teresa M.

APPLICANT: Lodes, Michael J.
 APPLICANT: Kalos, Michael D.
 APPLICANT: McNeill, Patricia D.
 APPLICANT: Veddyck, Thomas S.
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS
 FILE REFERENCE: OF HER-2/NEU-ASSOCIATED MALIGNANCIES
 CURRENT APPLICATION NUMBER: 210121.344
 CURRENT FILING DATE: US/09/930.125
 NUMBER OF SEQ ID NOS: 25
 SOFTWARE: FASTSEQ for Windows Version 3.0
 SEQ ID NO 1
 LENGTH: 3768
 TYPE: DNA
 ORGANISM: Homo sapien
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (1)...(3765)
 US-09-930-125-1

Alignment Scores:

Score:	1.16e-39	Length:	3768
Percent Similarity:	474.00	Matches:	261
Best Local Similarity:	36.85%	Conservative:	116
Query Match:	25.51%	Mismatches:	337
DB:	9	Indels:	312
		Gaps:	46

US-09-397-967-16 (1-1099) x US-09-930-125-1 (1-3768)

```

OY 223 ProCysGlyArg---LeuProGlyArgProTy---AlaLeuMetAlaLysTrIleLeu 240
DB 381 CCTGTGACAGAGGGCTCCAGAGAGCTGCGGAGCTGAGCTTCAACCCCTCACAGA 440
OY 241 AspLeuGlyArg---LeuHisProAlaIleThrIleuThrPheArgValIleu 258
DB 441 GATCTTGAAAGAGGGCTTGTATCCAGCG-----
OY 259 ProGlyAlaGlnGluProGlyLeuLeuArgValAlaGlyAspAsnGlyIleProTrp 278
DB 471 -----GAMCCCCACGCTGCTA-----CCAGA 494
OY 279 SerSerAsnAspGluLeuPheGlnThrPheCysAspPheProGlyIleValAspValSer 298
DB 495 CAC-----GATTTTGTGGA-----
OY 299 IleAsnGlnAlaProArgValGlyProAlaGlyIleuHisArgIleuValIleThrArg 318
DB 510 ---GGACATCTTCCACAGAACACAGCTGCTCTCAC-----
OY 319 MetAspGlyHis-----
DB 546 ACTGATAGACACCAACGCTCTCGGCTGCGACCCCTGTTCTCCAGTGTAAAGGCTTC 605
OY 323 ---IleLeuGlnAlaGluPheProGlyLeuProGlyAlaLeuSerPheValAlaLeuVal 341
DB 606 CCGCTGGTGGGAGAGAGCTTGTAGGATTCAGAG--CTGACG-----
OY 342 AspGlyIleuPheArgIleuIleCysAspSerArgHisIleThrPheCysIleuValAlaPro 361
DB 649 -----CGCACTGTCTGTGCGGCTGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG 696
OY 362 ProArgLeuLeuGlnGlu-----
DB 697 ACTGACGTGCGCATGACAGCTGCTGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 756
OY 369 -----AlaAspValCys-----HisGlyProIleThrIleu 378
DB 757 CTGGCCCTGCTCCATTCAACACAGCTGCACTGTGACCTGACCTGCTGCTGCTGCTGCT 816
OY 379 AspPheIleuIleHisIleuValAlaIleuValSerIleuProGlyIleuValIleuVal 398
DB 817 ACCTACACACAGACACGCTTTGATGTCATGCTCCATTCGCGAGGCGCGGAT----- 867
  
```

```

OY 399 ArgSerProGlnAspTrpAspSerPheLeuLeuThrAlaCysValGlnThrProLeuGly 418
DB 868 -----ACATGGGCGCCAGCTGTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 915
OY 419 ProAspTrpLysGlyCysLeuIleArgGlnAspProSerGlyAlaPheSerLeuValGly 438
DB 916 ACGAGCGTGGGATCTGCAACCT--CGTCTGCCCT-----
OY 439 LeuSerGlnProHisArgSerLeuArgGluLeuLeuAlaIleCysTrpAsp--SerGly 458
DB 951 ---GCACAACCAAGAGTGACAGAGGA-----TGGACACAGCGGTC 992
OY 458 euArg-----ValAspGlyAlaIleuValIleuValIleuValIleuValIleuVal 471
DB 993 TGAGAGTGCAGACAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1052
OY 472 -----AlaProArgProLysGluLysSerAsnLeuIleValAlaArg- 486
DB 1053 AGAGTGAGGCGCATGTACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1109
OY 487 -----GlyCysAsnProAlaProAla 494
DB 1110 CTTGGAGCTGCGCATTTCTGCGGAGAGCTTTGATGGGAGCCGCTCCAGACACTGC 1169
OY 494 roGlyCysSerProSer-----CysCysAlaLeu-----ThrGlnLeuSer 508
DB 1170 CCGGCTTCACGACAGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1229
OY 508 he-His-----ThrIlePro-----
DB 1230 ATACATCTCAGATGGCCGAGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1289
OY 513 -----ThrAspSerLeuGlnTrp-----His 519
DB 1290 AATCCGGAGACGATTTCTGCAATGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1349
OY 520 GluAsn-LeuGlnHisGlySerPheThrIlePheArgGlySerArgValGlyVal 539
DB 1350 CAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1403
OY 539 IAspGlyIleuThrHisAspSerGluValLeuLeuLysValMetAspSerArgHisArg 559
DB 1404 CCACCATTAACACCCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1463
OY 559 nCysMetGluSerPheLeuGlnAlaIleSerLeuMetSerGlnValSerTrpProHis 579
DB 1464 C-----CCGACCA 1472
OY 579 uValLeuLeuHisGlyValCysMetAlaGlyAspSerIleMetValGlnIlePheValTy 599
DB 1473 AGCTGTGCTGCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1517
OY 599 rLeuGlyAlaIleAspMetTrpLeuArgLysArgGlyHisLeuValSerIleSerTrp 619
DB 1518 CAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1565
OY 619 sLeuGlnValThrLys-----GlnLeuAlaIleValAlaLeuAsnTrpLeu 634
DB 1566 TCCAGGCGCCACCCAGTGTCACTGCAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1625
OY 634 uAsp-----LysGlyLeuProHisGlyAsnValSerAlaArgValValLeu 650
DB 1626 GGAATGCGAGTACTGCTGAGGCGCTCCCAAGGAGTGTGTGATGCCAGGACCTGTTGCC 1685
OY 650 uAlaArg-----GluGlyGlyAsp-----
DB 1686 GTGCAACCTGATGTGACAGCCCGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1745
OY 657 -----GlyAsnProPheIleIleLysLeuSerAspPro-- 667
DB 1746 CAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1805
  
```

```

QY 668 -GlyValSerProThrValIleuSerLeuGluMetLeuThrAspArgIleProTyrValAl 687
    ||||| ||| : : : : :
Db 1806 CAGTGGAAACCTGACCTCTCTACATGCCATC-----TGAAGTT 1847
QY 687 apProGlu-----CysLeuGlnGluAlaGlnThrLeuCysLeuG1 700
    ||| : : : : :
Db 1848 TCCAGATGAGGAGGGCGATGCCAGCTTGGCCATCACTGACACCCACTCTCTGTGTGA 1907
QY 700 uAlaAspLysTrpGlyPheGlyAlaThrThrTrpGluValPheGlnArgGlyProAlaH1 720
    ||| : : : : :
Db 1908 CCTGGATGACAGAGGCTGCCGCC-----GAGCAGAGAGCCAGCC 1949
QY 720 sIleThrSerLeuGluProAlaLysLysLeuLysPheTrpGluAspGlnGlyGlnLeuPr 740
    ||||| : : : : :
Db 1950 TCTGACGTCATCATCTCTGCGGTG-----GTTGGCATCTCTGT 1988
QY 740 oAlaLeuLysTrpThrGluLeuAlaGlyLeuIleThrGlnCysMetAlaTyrAspProG1 760
    ||||| : : : : :
Db 1989 GGTGCTGCTTGGGGTGGTCTTGGGATCTCATCAAG---CGACGGCAGCAGAGAT 2045
QY 760 YArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIleThrSerAspTy 780
    ||| : : : : :
Db 2046 CCGAGATGACAGATGCGGAGACTGCTGCAGAAAGC----- 2082
QY 780 rGluLeuLeuSerAspProThrProGlyIleProSerProArgAspGluLeuCysValAl 800
    ||||| : : : : :
Db 2083 -GAGCTGTGTGAGCGCGTGCACCTACGAGGATGCCACACAG----- 2127
QY 800 aGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGlnGluArgHisLeuLysTy 820
    ||||| : : : : :
Db 2128 ---CGCAGAGATG-----CGATCTCGAAGAGAGAGAGCTGAGAA 2165
QY 820 rIleSerLeuGlnGlyLysGlyAsnPheGlySerValGluLeuCysArgTyrAspProLe 840
    ||||| : : : : :
Db 2166 GGTGAAGGTCTGTGATCTGGCGCTTGGCAGATCTACAGGCACTGTGGATCCTGA 2225
QY 840 uGlyAspAsnThrGlyProLeuValAlaValLysGlnLeu---GlnHisSerValProAs 859
    ||||| : : : : :
Db 2226 TCGGAGAGATGTGAAATTCAGTGGCCATCAAGCTTGGAGGAGAAACACATCCCCAA 2285
QY 859 pGlnGlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuHisSerAspPheI1 879
    ||||| : : : : :
Db 2286 AGCCAAACAAGAAATCTTGACAGACATACGTGATGGCTGTGGCTCCCATATGT 2345
QY 879 eValLysTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuArgLeuValMetG1 899
    ||||| : : : : :
Db 2346 CTCCCGCTTCTGGGCATCTGCCG-----ACATCCACGGTGCAGCTGTGACACA 2396
QY 899 uTyrLeuProSerGlyCysLeuArgAspLeuGlnArgHisArgGly---LeuHisTh 918
    ||||| : : : : :
Db 2397 GCTTATGCTCTATGGCTGCTTACCATGTCGCGGAAACCCGCGACGCTGGGCTC 2456
QY 918 rAspArgLeuLeuLeuPheAlaTrpGlnIleCysLysGlyMetGluTyrLeuGlyAlaAr 938
    ||||| : : : : :
Db 2457 CCAGGACTCTGTAACGTGTATGAGATGCGAGGGATGAGCTACCTGAGAGATCT 2516
QY 938 gArgCysValHisArgAspLeuAlaIaArgAsnIleLeuValGluSerGluAlaHisVa 958
    ||||| : : : : :
Db 2517 GCGGCTGCTACACAGGAGCTTGGCGCTCGAAGCTGTCTGTCAAGAGTCCCAACCATGT 2576
QY 958 lLysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAspTyrTyrValVa 978
    ||||| : : : : :
Db 2577 CAATAATTACAGACTTGGGGGTGGCTGCGTGTGACATGACAGAGACAGATGACATGC 2636
QY 978 lArgGluProGlyLysSerProIlePheTrpTyrAlaProGluSerLeuSerAspAsnI1 998
    ||||| : : : : :
Db 2637 A---GATGGGGGAGAGGCGCATCAAGATGAGCCCTGAGTCCATCTCCGCGGCG 2693
QY 998 ePheSerArgGlnSerAspValTrpSerPheGlyValValleuTyrGluLeuPheThrTy 1018
    ||||| : : : : :
Db 2694 GTTCACCCACAGAGATGATGTGTGAGTATGTGTGACTGTGTGGAGCTGATGACTT 2753
QY 1018 rCysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetGlyProGluArgG1 1038

```

```

Db 2754 TGGGGCCAAA-----CCTACGA 2771
QY 1038 uGlyProPheLeuCysArgLeuLeuGluLeuAlaGlnGlyArgGluProProPr 1058
    ||||| : : : : :
Db 2772 TGGGATCCCAAGCCCGGAGATCCCTGACCTGTGAAAAGGGGAGCGGCTGCCAGCC 2831
QY 1058 oProThrCysProThrGluValGlnGluMetGlnLeuCysTrpAlaProGluProH1 1078
    ||||| : : : : :
Db 2832 CCCCATCTGCACCATGATGATCTACATGATCATGTGTCAATGTGTGATGATGACTCTGA 2891
QY 1078 sAspArgProAlaPheAlaThrLeuSerProGlnLeuAspProLeuTTPArgGlyArgPr 1098
    ||||| : : : : :
Db 2892 ATGTGGCCAAAGATTCGGGAGTGTGTGTGAT--TCTCCGATGAGCCAGGAGACCC 2949
QY 1098 o 1098
Db 2950 C 2950

Search completed: April 28, 2003, 18:58:40
Job time : 419.323 secs

```

GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 28, 2003, 13:51:29 ; Search time 150.669 Seconds

(without alignments)
2236.936 Million cell updates/sec

Title: US-09-397-967-16

Perfect score: 5860

Sequence: 1 MAPSEETPLIFQRCSSLS.....RPAPATLSQDPLMRGPG 1099

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 441362 segs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_{p2n}.model -DEV=xlp
-O=/cgn2_1/USPTO_spool/US09397967/runat.26042003.182313.10750/app.query.fasta.1.1486
-DB=Issued_Patents_NA -GFWT=fastcap -SUFFIX=rnt -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=trans40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pro -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09397967 -ECGN_1_1=59.6runat.26042003.182313.10750 -NCPU=6 -ICPU=3
-NO_XIPXY -NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: Issued_Patents_NA:*

- 1: /cgn2_6/plodata/2/1na/5A.COMB.seq:*
- 2: /cgn2_6/plodata/2/1na/5B.COMB.seq:*
- 3: /cgn2_6/plodata/2/1na/6A.COMB.seq:*
- 4: /cgn2_6/plodata/2/1na/6B.COMB.seq:*
- 5: /cgn2_6/plodata/2/1na/PCUS.COMB.seq:*
- 6: /cgn2_6/plodata/2/1na/backfillseq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5223	89.1	4016	5	PCT-US95-08354A-1
2	4418	75.4	3807	1	US-08-357-598-1
3	4418	75.4	3807	2	US-09-003-289-1
4	4418	75.4	3807	5	PCT-US95-16435-1
5	2629	44.9	3435	4	US-09-046-158A-21
6	2621	44.7	4482	3	US-08-567-508C-1
7	2621	44.7	4482	3	US-09-196-480-1
8	2619.5	44.7	3629	3	US-08-097-997A-8
9	2619.5	44.7	3629	3	US-08-665-574C-8
10	2619.5	44.7	3629	3	US-08-946-994-8
11	2360	40.3	3495	1	US-08-446-038B-2
12	2360	40.3	3495	1	US-08-446-010B-2

13	2360	40.3	3495	1	US-08-805-445-2
14	2360	40.3	3495	2	US-08-064-067D-2
15	2360	40.3	3495	2	US-09-066-208-2
16	2360	40.3	3495	4	US-08-980-080-3
17	1904	32.5	3429	1	US-08-097-997A-10
18	1904	32.5	3429	3	US-08-665-574C-10
19	1904	32.5	3429	4	US-08-946-994-10
20	1904	32.5	4234	1	US-08-446-038B-1
21	1904	32.5	4234	1	US-08-446-010B-1
22	1904	32.5	4234	1	US-08-805-445-1
23	1904	32.5	4234	2	US-08-064-067D-1
24	1904	32.5	4234	2	US-09-066-208-1
25	1815.5	31.0	3561	1	US-08-097-997A-12
26	1815.5	31.0	3561	3	US-08-665-574C-12
27	1815.5	31.0	3561	4	US-08-946-994-12
28	482	8.2	3845	2	US-08-220-240A-4
29	475	8.1	4530	1	US-08-229-515A-9
30	475	8.1	4530	1	US-08-645-865-9
31	475	8.1	4530	1	US-09-167-332-4
32	474	8.1	3768	2	US-08-625-101-1
33	474	8.1	3768	2	US-08-356-786-1
34	474	8.1	4473	2	US-09-048-804-1
35	474	8.1	4473	4	US-09-056-105-26
36	466.5	8.0	5532	2	US-08-475-035-3
37	466.5	8.0	5532	4	US-09-676-610B-17
38	465.5	7.9	5484	4	US-09-632-580A-3
39	465.5	7.9	5501	1	US-08-484-438-1
40	465.5	7.9	5555	1	US-08-484-438-3
41	465	7.9	8083	4	US-09-383-630-4
42	465	7.9	8083	4	US-09-383-630-5
43	464.5	7.9	3955	1	US-08-229-515A-14
44	464.5	7.9	3955	1	US-08-645-865-14
45	462	7.9	4508	5	PCT-US93-06251-34

ALIGNMENTS

RESULT 1
PCT-US95-08354A-1
Sequence 1, Application PC/TUS9508354A
GENERAL INFORMATION:
APPLICANT: Temple University - Of The
APPLICANT: Commonwealth System of Higher Education
TITLE OF INVENTION: JAK3 PROTEIN TYROSINE
TITLE OF INVENTION: KINASE AND DNA ENCODING THE SAME
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seidel, Gonda, Lavoorgna
ADDRESSEE: 6 Monaco, P.C.
STREET: Suite 1800, Two Penn Center
CITY: Philadelphia
STATE: Pennsylvania
COUNTRY: U.S.A.
ZIP: 19102
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 720 kb
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/08354A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/232,368
FILING DATE: 8 July 1994
ATTORNEY/AGENT INFORMATION:
NAME: Monaco, Daniel A.
REGISTRATION NUMBER: 30,480
REFERENCE/DOCKET NUMBER: 6056-203 PC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-8383
TELEFAX: (215) 568-5549

; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 4016 nucleotides
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single stranded
 ; TOPOLOGY: linear
 ; PCT-US95-08354A-1

Alignment Scores:

Pred. No.:	0	Length:	4016
Score:	5223.00	Matches:	1060
Percent Similarity:	80.59%	Conservative:	3
Best Local Similarity:	80.36%	Mismatches:	32
Query Match:	89.13%	Indels:	235
	5	Gaps:	6

US-09-397-967-16 (1-1099) x PCT-US95-08354A-1 (1-4016)

```

QY 1 MetAlaProSerSerGluThrProLeuIleProGlnArgSerCysSerLeuSerSer 20
    |||
Db 7 ATGGCACCCTCCAGTAGAGACACCTCTGATCCTCAGCGCTCTTACAGCTTCATCC 66
    |||
QY 21 SerGlnAglAlaLeuHisValLeuLeuProProArgGlyProGlyProProGlnArg 40
    |||
Db 67 TCAGAGGAGAGAGCCCTGCATGTCTCTTCCCTCCGGGAGCTGGGCTCCCGACGGA 126
    |||
QY 41 LeuSerPheSerPheGlyAspTyrLeuAlaGlnAspLeuValArgAlaAlaValAla 60
    |||
Db 127 TTGCATCTCTCTTTGGGAGTACTGCTGAGATTTATGTGGCAGCTCCCAAGGCC 186
    |||
QY 61 CysGlyIleLeuProValTyrHisSerLeuPheAlaLeuAlaThrGlnAspPheSerCys 80
    |||
Db 187 TGTGCGATCCGCTGCTTTATCATCTGCTTTGCTGCGCTGAGACTTCCTTCC 246
    |||
QY 81 TrpPheProSerHisIlePheCysIleGlnAspValAspThrGlnValLeuValTyr 100
    |||
Db 247 TGGTTTCCCAAGCCACATCTCTGATAGAGAGCGTGCACATCAAGTCTTGCTTAC 306
    |||
QY 101 ArgLeuArgPheTyrPheProAspTyrPheGlyLeuGlnThrCysHisArgPheGlyLeu 120
    |||
Db 307 AGGCTACGCTTTATTTCTCTGACTGGTTGGCTGGAGACATGTACACCGCTTTGGGCTG 366
    |||
QY 121 ArgGlyAspLeuThrSerAlaIleLeuAspLeuHisValLeuGlnHisLeuPheAlaGln 140
    |||
Db 367 CGCAAGATTTGACACAGTGCATCTTACTTACTTATAGAACATCTTTGCTTCAG 426
    |||
QY 141 HisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetGlyGlnGlnGly 160
    |||
Db 427 CACCCAGTGCAGTGTAGTGGGCTCCCGGTGGGCTTTAGCATAGAGAGAGAGGGA 486
    |||
QY 161 GluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGlnAlaGlnArg 180
    |||
Db 487 GAGTTCCTGAGCGCTGGCTGTGCACTTGGCCAGATGGCTCGAGACAGCCAGCGC 546
    |||
QY 181 ProGlyGlnLeuLeuLysThrValSerTyrLysAlaCysLeuProProSerLeuArgasp 200
    |||
Db 547 CCAGAGAGACTGCTGAACACGCTCAGTTACAAAGCCTGTCCGCCACAGCTGGCCAT 606
    |||
QY 201 ValIleGlnGlyGlnAspPheValThrArgArgArgIleArgArgThrValValLeuAla 220
    |||
Db 607 GTGATCCAGGGCCAGAACTTCTGTGACAGCAGCGCCATCCGAGAGACGCTGTGGCG 666
    |||
QY 221 LeuLeuProCysGlyArgLeuProGlyArgProTyrAlaLeuMetAlaLysTyrIleLeu 240
    |||
Db 667 CTCGC--GGCTGTGGTGGCTGCGCAGCGCAGCGCTACAGGCTCATGGCAAGATAT-CTG 723
    |||
QY 241 AspLeuGlnArgLeuHisProAlaAlaThrThrGlnThrPheArgValGlyLeuProGly 260
    |||
Db 724 GACCTGAGACGGCTACATCCAGCGCCACACAGACCTTCCGTGGGGCTCCCGGGC 783
    |||
QY 261 AlaGlnGlnGlnProGlyLeuLeuArgValAlaGlyAspAsnGlyIleProTyrSerSer 280
    |||
Db 784 GCCCAGGAGAGCCGGGCTTCTGCTGTGGGGGAGACAGCGCATCTCTCTGGAGCTCC 843
  
```

```

QY 281 AsnAsp----- 282
    |||
Db 844 GGGAGCCAGAGAGTGTGGGGCTGGGCTTGAGAAAGCGGGGTCCGGGGAGGCGTGGCGG 903
    |||
QY 282 ----- 282
    |||
Db 904 TTAGTAAGAAAGCCGTCGGAACAATCAGAAAGTGCTAGTCCGGGGCGGCTGATTGC 963
    |||
QY 282 ----- 282
    |||
Db 964 ACTACGGGGCAGGGGGGGGGGTCAATTTAAAGTGGTCCGGGAGTGGAGCTGCCCGG 1023
    |||
QY 282 ----- 282
    |||
Db 1024 GGACTGACCTGGGGCGTGAAGTGAAGGTGAGACTAGTAGGGGGCGGGAGCAGAGAA 1083
    |||
QY 282 ----- 282
    |||
Db 1084 GCGGATAGTATGATTGATCAGTGGGCATGACCCGACAGGCTGTCTTCACTCCCTTA 1143
    |||
QY 282 ----- 282
    |||
Db 1144 ACCTGTCCTCCCCCCCCCAGCAGTGGAGGTGGGAGAGAGAGAGAGAGAGAGTGTGCT 1203
    |||
QY 282 ----- 282
    |||
Db 1204 GGGTCAGTCTGTAAATCCCTGCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1263
    |||
QY 282 ----- 282
    |||
Db 1264 TTTGCGAGACAGGATTCTCTGATAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1323
    |||
QY 282 ----- 282
    |||
Db 1324 GGCTGGCTGGAACCTGAAATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1383
    |||
QY 283 ----- 283
    |||
Db 1384 CGTGTGCCACTACGCGCTACTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1443
    |||
QY 290 AspPheProGlnIleValIleAspValSerIleAsnGlnAlaProArgValGlyProAlaGly 309
    |||
Db 1444 GACTTTCGGAATACGTGATGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1501
    |||
QY 310 GluHisArgLeuValThrValThrArgMetAspGlyHisIleLeuGlnAlaGlnPhePro 329
    |||
Db 1502 AG-CACCGGCTGTACCTGTCACCCAGAGTGAACGCCATCCTGGAGAGCGAGTTCCG 1560
    |||
QY 330 GlnLeuProGlnAlaLeuSerPheValAlaLeuValAspGlyTyrPheArgLeuIleCys 349
    |||
Db 1561 GGGCTGCTGAGGGCGCTGTCTTTGCTGGCCCTCGTGGATGGTACTTCGCCCTGATCTGC 1620
    |||
QY 350 AspSerArgHisTyrPheCysLysGlnValAlaProProArgLeuLeuGlnGlnAla 369
    |||
Db 1621 GACTCCAGGACTTATTTCTGCAAGAGAGTGGCGCCAGCGCTGTGAGAGAGAGCGC 1680
    |||
QY 370 AspValCysHisGlyProIleThr----- 377
    |||
Db 1681 GAGCTGTGCCATGAGACCATCAGCTTAGAGAGCTAGTGGGTCAACAGTGGGAGTGGGGA 1740
    |||
QY 378 ----- 380
    |||
Db 1741 TGGCAGCGAGAGCTGTCTTGGTTGGACCTGGCTGTCTTGGCTAGGTTAGACTTT 1800
    |||
QY 381 AlaIleHisLysLeuLysAlaAlaGlySerLeuProGlyThrTyrIleLeuArgArgSer 400
    |||
Db 1801 GCCATCCAGCAGTGAAGAGCGCGTGGCGTCCAGGACCTATATCTCCCGCCGACG 1860
    |||
QY 401 ProGlnAspTyrAspSerPheLeuLeuThrAlaCysValGlnThrProLeuGlyProAsp 420
    |||
Db 1861 CCGCAGGACTATGACAGCTTCTTTTACCGCTGCGCTGAGACTCTCTTGGGCCCGGAC 1920
  
```


Oy	421	TtyysgIyCysLeuIleArgGInAspProSerGlyAlaPheSerLeuValIleuSer	440
Dd	1921	TACAAAGGCGCTCCCTCATTCCCGCAGAGCCACGAGGGGCTTTCTCCCTGGCTCC-CTCAAC	1979
Oy	441	GInProHisArgSerLeuArgGluLeuLeuAlaAlaCysTrpAsnSerGlyLeuArgVal	460
Dd	1980	CAGCCCCACAGAAACCTCGGGAGAGCTGGCTTCAGC-TGGCTGAATTCC-GGGCTGGAGATA	2037
Oy	461	AspGlyAlaAlaLeuTrpLeuThrSerCysGlyAlaProArgProLysIleuSerAsn	480
Dd	2038	GAGCGTGTGGCTGAACCTTAACATCTGTGGCTGCCAGACCAGAAAGAAAGTCCAT	2097
Oy	481	LeuIleValValArgArgGlyCysAsnProAlaProAlaProGlyCysSerProSerCys	500
Dd	2098	TTGATATCGTGGTGGCAAGGGGCTGACCCCCGGCGCTGCCCTGGCTGCCCTCCGCTCCG	2157
Oy	501	CysAlaLeuThrGlnLeuSerPheHisThrIleProThrAspSerLeuGluTrpHisGlu	520
Dd	2158	TGTCCGCTGACACAGCTGAGACTTCCACACAAATTCACAAGGACAGCGTGG---GGACACAG	2214
Oy	521	AsnLeuGlnHisArgSerPheThrLysIlePheArgGlySerArgArgGluValAlaAsp	540
Dd	2215	AACCTGGGTACGGTCTTTTACCAGATCTTCGTGGCGCCAGCGGGAGGTCTGGAT	2274
Oy	541	GlyGluThrHisAspSerGluValLeuLeuLysValMetAspSerArgHisArgAsnCys	560
Dd	2275	GGTAGAGACACTGTACTGGAAAGTCTCTGTAAAGTATGAGATCCAGATCGGAACATGTC	2334
Oy	561	MetGluSerPheLeuGluValAlaSerLeuMetSerGluValSerTrpProHisLeuVal	580
Dd	2335	ATGAGACTTTTCTCGAAGCCCGCAAGCTTGATATGACCAAGTATCTTACCCTGGACCTGGTG	2394
Oy	581	LeuLeuHisArgLysValCysMetAlaGlyAspSerIleMetValGlnIlePheValTrpLeu	600
Dd	2395	TTACTGACAGCGCGTCTCATGGCTGGAGACAGCATCATGTGCAGGAATTTGTATCTTA	2454
Oy	601	GlyValIleAspMetGlyLeuArgLysArgGlyHisLeuValSerTrpAsnTrpLysLeu	620
Dd	2455	GGAGCAATTGCATGTACTGCTGGCCAGCGGAGCTGGCCACTGGTGTACACCAGCTGGAACTG	2514
Oy	621	GlnValThrLysGlnLeuAlaTrpAlaLeuAsnTrpLeuGluAspLysGlyLeuProHis	640
Dd	2515	CAGGTGACCAAGCAGCTGGCATATGACCTTAATCTATTGGAGAGCAAAAGGCTTCTCAGC	2574
Oy	641	GlyAsnValSerAlaArgLysValLeuLeuAlaAlaArgGlnGlyLysAspLysAsnPro	660
Dd	2575	GGCAACGCTTACGACGGAAGAGGTCTCTGGCTGGAGGGGGGTATGGAAATCCACT	2634
Oy	661	PheIleLysLeuSerAspProGlyValSerProThrValLeuSerLeuGluMetLeuThr	680
Dd	2635	TTTCATTAAAGCTGATGATCCTGGGTGATGATGCCACATGCTGTACGCTGGAAATGCTCAC	2694
Oy	681	AspArgIleProTrpValAlaProGlyCysLeuGlnGluAlaGlnThrLeuCysLeuGlu	700
Dd	2695	GACAGAAATACCTCTGGGTGGCCCCCGAATGTCTCCAGAGAGCTCAGACACTGCTTGGAG	2754
Oy	701	AlaAspLysTrpGlyPheGlyAlaThrThrTrpGluValPheGlnArgIleProAlaHis	720
Dd	2755	GCTGACAGTAGTGGGCTTTGGAGCCACCACCTGGAGAGGTGTTCAGCGGGGAGACCGCCAC	2814
Oy	721	IleThrSerLeuGlnProAlaLysLysLeuLysPheTrpGluAspGlnGlyLeuPro	740
Dd	2815	ATCACCTCGTGGTGGAGCCCGCAAAAGCTGAATGTTATGAGGACCAAGGACAGCTGCC	2874
Oy	741	AlaLeuLysTrpThrGlnLeuAlaGlyLeuIleThrGlnCysMetAlaLysAspProGly	760
Dd	2875	GCTCTCAATATGACAGAAAGCGGGGCTTATACACAGAGCATGGCGTATGATCTGTGC	2934
Oy	761	ArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIleHisSerAspTrp	780
Dd	2935	CGGCGCCGCTCTTCCGAGCATCTCTAGAGACTTACAGCGCTCTATTACATCAAGATTTC	2994
Oy	781	GluLeuLeuSerAspProThrProGlyIleProSerProArgAspGluLeuCysValAla	800

```

Db      2995 GAGCTCTCCA -GACCCCAACTGGCATTCGGAGTCTTCGAGATGAGCTGTGCCT --- 3049
Oy      801 G1YAlaGLInLeuTyraLaCySGInAsPProIaLlePhneGIuGLuaRHisLeuLyThr 820
Db      3050 GCGGCCAGCTCTATTGCTCCTCCAGAGACC GCCCATTTTC -GAGGAGAGAACCTTAAGTAC 3108
Oy      821 IleSerLeuLeuGLysGLysAsnPhneGLyservaGLuLeuCysArgTYrAspProleu 840
Db      3109 ATCTCTTGCTGGGGCAAGGGCAACTTTGGAGCGTAGGACTGTGGCTGGCTATGACCCCTTG 3168
Oy      841 GLYASpaSnThRGlyProLeuValAlaValLySGInLeuGLNHisSerVaLProasPLn 860
Db      3169 -- -GACAATACGGGACCCCTGGTGGCAGTGAAACACACTACACAGCGGGCCAGACAG 3225
Oy      861 GlnArgaspPhneGLInatrgLIuIIeGLInLeuLYSAlaLeuHisSeraspHeIIevaI 880
Db      3226 CAGAGGGCACTTCACGCGGGAATTCAGATCTTTAAGGCTCTCACAGCACTTCATCTTC 3285
Oy      881 LysTYrAtRGlyALysaTYrGLYTrgLYProGLYArgGSInSerLeuArgLeuValMetGLuYr 900
Db      3286 AAGTACCGGGAGTGCACACTATTGGGCCAGAGTGGCCAGAGCCCTGCCTGTGTATGAGATAC 3345
Oy      901 LeuProSerGLYCYSLeuArgaspLeuArgaspLeuGLInatrgHisArg -- GlyLeuHisThrAsp 919
Db      3346 CTGGCCACAGCGGCTTCCTCGCAGAGACTTCCTGCACGCCCATCGCGCGGCTTCGACACCCAG 3405
Oy      920 ArgLeuLeuLeuPhenalaTrrpgInIIECYLSYGImetGLuTYrLeuGLYAlaArgAr 939
Db      3406 CGCCTACTGCTGTTCGCTTGGCAGATCTGCAGAGGATGAGATGAGTGGGTGGCGGCGCG 3465
Oy      940 CysValHISArgaspLeuAlaAlaArgAsnIIELeuValGLSerGLYAlaHISValLYs 959
Db      3466 TGGCGTACACCTGTGACCTGGCTGGCGGCAACATTTGTGTGAAGAGGAGCTCATGTATGAAG 3525
Oy      960 ILEAlaSPheGLYLeuAlaILysLeuLeuProLeuGLYLSASPtyTYrTYrAlaVArg 979
Db      3526 ATCCGGACATTCGGGCTTCGCTAGCTGCTGCCCCCTGGGAAGACTACTACTGGTCCGC 3585
Oy      980 GluProGLYInSerProIlephETPYrAlaProasLeuSerLeuSeraspnaIllephe 999
Db      3586 GAGCGTGCGCAAACCCCACATCTTTGGTATGCCCGGAGTCCCTATCTGACAAACATCTTC 3645
Oy      1000 SerArgInseraspValITrpSerPhneGLYValValLeuTYrGLuLeuPhenhetrycys 1019
Db      3646 TCCCGCAATCTGACGTGAGACTTCGGAGTGTGTATTACAGACTCTTCACCTACCTGC 3705
Oy      1020 AspLysSerCYSSerProSerAlaGLuPhLeuArgMeIeMGlyProGLuArgGLuLY 1039
Db      3706 GACAAGAAGCTGCACACCATCCGCGAGATTCCTGGCATATGGGGCTGTAGAGCTTAAGCA 3765
Oy      1040 ProProLeuCYsATrLeuLeuGLuLeuLeuAlaGLuGLYArgArgLeuProRobroPro 1059
Db      3766 CCCCCGCTGTCGCCCTCTGTGAGACTCTGGCGAGAGGGCGGAGCGCTCCACACACTGCC 3825
Oy      1060 ThrCysProThrgInuValaGLInGLuLeuMetGLInLeuCYSTrPaLaProGLuProHisasp 1079
Db      3826 ACCTGCCCAACCGAGGTTCAGAGACTCATGCACTGTGGCTGGCGCCAG -CCGACAGAC 3884
Oy      1080 ArgProalaphelaTrhrLeuSerProGLInLeuAspProLeuTYrArgGLYArgPro 1098
Db      3885 CGGCGAGGCTTCGGCAGCCCTGAGAGCCCGCACACTGGAGGCGCTGNGGGGTGAAGACCC 3941

RESULT 2
US-08-357-598-1
; Sequence 1, Application US/08357598
; Patent No. 5705625
; GENERAL INFORMATION:
; APPLICANT: Clivin, Curt I.
; APPLICANT: Small, Donald
; TITLE OF INVENTION: NOVEL PROTEIN TYROSINE KINASE, JAK3
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:

```

Tue Apr 29 06:04:28 2003

ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/357,558
FILING DATE: 15-DEC-1994
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: Haile, Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07265/033001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3807 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: DNA

IS-08-357-598-1

Alignment Scores:	
pred. No.:	0
Score:	4418.00
Percent Similarity:	86.94%
Best Local Similarity:	81.16%
Query Match:	75.39%
DB:	1
US-09-397-967-16 (1-1099) × US-08-357-598-1 (1-3807)	
	length: 3807
	Matches: 870
	Conservative: 62
	Mismatches: 130
	Indels: 12
	Gaps: 6

Qy	1	MetLaPrProSerGluInlThPLeuLleProGlnAmgSerCysSerLeuSer	20
Db	168	ATGGCACTCCAACTGAAGAGACCCCTGATCCCTACGGTTATACAGCTCTTCTTC	227
Qy	21	SerGluLaGlyAlaLeuHisValLeuLeuProGlnGlyProGlyProGlnArg	40
Db	228	ACGAGAGCTGTGCCCTGCATCTCTGTGCGCCCTGGGGCCCCCGCCACGC	287
Qy	41	LeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysValArgAlaAlaLysAla	60
Db	288	CTATCTTTCTCTTTGGGGACCACTGGCTGGAGACCTGTGCGAGCTGCCAACGCC	347
Qy	61	CysGlyLleLeuProValTyrHisSerLeuPheAlaLeuAlaTrpGlnAspPheSerCys	80
Db	348	AGCCGATCTGCTGCTGTACCACTCCCTCTTCTCTGTGGCACGGAGACTCTCTGC	407
Qy	81	TrpPheProPheSerHisIlePheCysIleGluAspValAspThr-GlnValLeuValTyr	100
Db	408	TGGTT-CCCCGAGACCACTTTCTCCGTGGAGGATGCACACCCCAAGCTCGTGTGA	466
Qy	100	ArgLeuAlaArgPheTyrPheProAspTrpPheGlyLeuGlnThrCysHisArgPheGlyLe	120
Db	467	CAGGATTCGCTTTACTTCCCAATTGGTTGGCTGGAGAAAGGCAACGCGCTTGCGGCT	526
Qy	120	uArgIAspSerPheThrSerAlaIleLeuAspLeuHisValLeuGlnHisLeuPheAlaG	140
Db	527	ACGCAAGATTGGCCAGTGCATTACCTTGACCTCCAGTCTCTGGAGACACTTTGGCCA	586
Qy	140	HisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLysGluInlG	160
Db	587	GCACGCAAGTACCTGTGATGGGCGGCTCCCGCGTGGCTCAGTGTCAAGAGAGCAGG	646

us-09-397-967-16.rni

Page 4

QY	160	YGNPHEUSENLEUVALALEUASPHEUVALAGLMERLALARGLUGLINLAGINAR	180
Db	647	TGAGTGTCTACGCTGGCGGCTGTTTGACCTTGGCCCCGGATGGCGGAGACGACGCCACGGC	706
QY	180	GPRLGLLEULEUENLYSTHVALSETTYLVSALACYSLEUPROBROSERLEUARGAS	200
Db	707	GCGGGAGACGCTGCTGAGAGCTGTACAGTACAGAGCCTGCTACCCCAAGCCTGGCGCA	766
QY	200	PVALLEGLINLYGINSNPHEVALTHIRARGARVILARGARGHTRVALLEUAI	220
QY	220	ALEUENPROCYSGLYARGLEUPROGLIARGPRO-TYRALALEMERLALYSTYRTEL	240
Db	827	CCTGGC-CGGGTGGCGGCTTGGCAGGAGACGCCGACATCGCTCATGGCCAAATCATCA	885
QY	240	EUASPHEUGLNUARGLEUHSIPROALAATHRTHTGUTHIRPHEARGVALGLYLEUPROG	260
Db	886	TGGACCTGGAGCGGTGGATCCACCGGGCGCGGACGACCTTCACAGTGGGCTCCCTG	945
QY	260	LYALA-----GlnGluGluProGlyLeuLeuArgValAlaGlyAspAsnGlyLeu	277
Db	946	GGGCCCTTGTTGGCCACGACGGGCTGGGGCTCTCCCGCTGGCTGGTGGAGGGGGCATCG	1000
QY	277	ROTTPSERASNAASP--GluLeuPheGlnThrPheCysAspPheProGluLeuValA	296
Db	1006	CCTGGACCCAGGAGAAACAGGAGGCTCCCTCCAGCCTTCTCGATTTCCGAATGCTAG	1060
QY	296	SPVALSERILASNGLINLAPROARGVALGLYPROVALGILGILNHSARGLEUVALTHRY	316
Db	1066	ACATTAGCATACAGCAGGCCCCCGGCTTGGCCCCGGCGGAGACACCGCTGGTCACTG	112
QY	316	ALTHIRARGMELASPOLYHISLEUGLNUAGLUPHEPROGLYLEUPROGLIUALAENS	336
Db	1126	TTACACGACGACAGAACACAGATTTTTAAAGCGCGAGTTCCAGAGGCTGCCAGAGCTCTGT	118
QY	336	ERPHVALALALEUVALASPOLYTYRPHARGLEULIECYSASPSEARHISTYRPHEC	356
Db	1186	CGTTGTGGCCCTGTGGACGCGTACTTCCGGCTGACACCGAGATCCCAACACTCTCT	124
QY	356	YSLYSGLIVAlAlAPROARGLEUENUGLUGLNUALASPVALCYSHISGLYPROI	376
Db	1246	GCAAGAGAGTG---GACCCGAGGCTGGTGGAGAAATGGCGGACACATGCCACAGGCCCA	130
QY	376	LETHRLEUASPHEALALEHISLYSLEUYSAALAGLYSERLEUPROGLYTHRYT	396
Db	1303	TCACCTCGAGCTTGGCCATCAACAGCTCAAGACTGGGGGTCTACGTCGCGCTCATG	136
QY	396	LELEUARGIRSERPROGLINSPHYRASPSERPHELEULEUETHALACYSVALGILTHRP	416
Db	1363	TYTCTCGCCGATCCCCAGACTTTGACACTTCTCTCTACTGTGTGTCCAGAAC	142
QY	416	ROLEUGLYPROASPIRYLDSGLCYSLYLEUARGGLNASPPOSERGILYALAPHESERL	436
Db	1423	CCCTTGGCTGATTAATTAAGGGCTGGCCATCGCGCCAGCGCCACAGAACTTCTTTC	148
QY	436	EUVALGILYLEUSERGLNPROHISARGSERLEUARGLILEULEUVALALACYSTIPASNS	456
Db	1483	TGGTTTGGCTTAGCCGACCCACACAGCATCTTCCAGAGCTCTGGCAACCTCGGGATG	154
QY	456	ERGLYLEUARGVALASPOLYALALALEUUTYLEUTNRSERYCYSVALARARGPROL	476
Db	1543	GGGGGCTGACAGTAGATGGGTGGCAGTGAACCCCTACCTCTCTGTTATGCCAGACCA	160
QY	476	YSGLILYSSERASNLEUILEVALVALATARGGLCYSPASNPROALAPROALAPROGLYC	496
Db	1603	AAGAAATCTCAACCTGATTTGTGTCCAGAGAGCTCAACGCCACCCAACTATCTTGG	166
QY	496	YSSERPRO---SERCYCYSALEUUTHTHTGILNHSERPHEHISETHRLIPROTHASPS	515
Db	1663	TTCAAGCCCCCATCCAAATACACAGCTGAAGTACATGATCAATTTCCAAAGAATCCCTGTGACA	172
QY	515	ERLEUGLUTTRHISGLIUNSNLEUGLYHISGLYSERPHERHLYSLIEPHARGGLYSERA	535

```

Db 1723 GCGGTGGAGTGCATGAGAACTGGGCCATGGGCTCTTCACCAAGATTATTCAGGGGCTGTG 1782
Oy 535 rgaatgvalvalaaspglvglturhrhlaaspsrgluvalleuylvalmetasp 555
Db 1783 GCGCATGAGTGGTGGATGGGAGGCCGGAAGACAGAGAGCTGTCTGTGAAGTCAATGATG 1842
Oy 555 eAtArhNlaAtgAsnCysMeTgluSerPheleuGluaAlaSerleuMetserGluaLs 575
Db 1843 CCAAGCAACAAGAACTGATGATGATTCCTGTGAAGGACGAGCTGTATGAGCCAAAGTGT 1902
Oy 575 eTyrPrroHlaSerValleuLeuunhlsGlyValCysMeAlaGlyAspserlleMetVal 595
Db 1903 CGTACCGGCAATCTGTGTGTCTCCACGCGCTGTGCATGGCTGGAGACACGCAATGTGCG 1962
Oy 595 lngluPheValTyrleuGlyAlaAlaIleAspMetTyrleuArglysArglylnHlaSerVal 615
Db 1963 AGGAATTTGTACACCTGGGGCCATAGACATGTATCTCGAANAAGCTGGCCACCTGGTGC 2022
Oy 615 eAlaSerTrpLysleuGlnValThrLysGlnleuAlaTyrAlaLeuAsnTyrleuGln 635
Db 2023 CAGCAGCTGGAACTGGAGTGGTCAACAGCTGGCTACGCTCAACTATCTGAGG 2082
Oy 635 ePlyrsglyLeuProhNlaGlyAsnValSerAlaArglyValleuLeuAlaArgGly 655
Db 2083 ACAAGGCTGTCTCCATGGCAATGTCTGTGCGGAGAGTCTGTGGGAGGAGG 2142
Oy 655 lyaSPglYAsnProPheIleLysleuSerAspProGlyValSerProThrValLeu 675
Db 2143 CTGATGGGAGAGCCGCTTCATCAAGCTGAGTGAACCTGGGGGTACGCCCTGTGTAA 2202
Oy 675 eTleuGlnMetleuThrAspArgTleProTrpValAlaProGlnCysleuGlnAlaG 695
Db 2203 GCGTGGAGATCTCACGACAGATCCCTGGGTGGGCCCGGAGTGTCTCGGGAGGCGC 2262
Oy 695 lntPrleuCysleuGlnAlaAspLysTrpGlyPheGlyAlaThrThrTrpGluValPheG 715
Db 2263 AGACACTTAAGTTGGAGTACAGAGTGGGCTTGGGCGCACGCTGCGGAGAGTGTAA 2322
Oy 715 lnaTgGlyProAlaHlaIleThrSerleuGlnProAlaLysleuLysPheTyrGln 735
Db 2323 GTGGCGTCACCATGCCCATGTGCTGCTAGATCTGCTGAAGAACTCAATTTTGTGAGG 2382
Oy 735 spGlnGlyGlnleuProAlaLeuLysTrpTrpGlnleuAlaGlyleuIleThrGlnCysM 755
Db 2383 ACCGCGACAGACTGTGCGGCCGCAAGTGGACAGAGCTGGCTGTATCAACAGTCCA 2442
Oy 755 eTalaTyrAspProGlyArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyL 775
Db 2443 TGGCTATAGACCGGCTCCAGAGGCTCTCTTACGAGCGCTGATTCGTACCTCAATAGTC 2502
Oy 775 euIleThrSerAspTyrGlnleuLeuSerAspProThrProGlyIleProSerProAla 795
Db 2503 TCATCTCTCTGACTATGAGTCTCTCTAGACACACAC---TGGTGGCTGGCACTGCTG 2559
Oy 795 spGluLeuCysValAlaGlyAlaGlnleuTyrAlaCysGlnAspProAlaIlePheGln 815
Db 2560 ATGGGCTGTGG---AATGGTCCCAAGCTCTATGCTTCCCAAGACCCCAAGCTCTTCCAGG 2616
Oy 815 lnaTgHlaIleLysleuTyrTleSerleuGlnGlyLysGlyAsnPheGlySerValGluLeu 835
Db 2617 AGACACACCTTAAGTACATCTCACAGCTGGGCAAGGCTCTTGGCAGCGGAGGCTGT 2676
Oy 835 ysaTgTyrAspProleuGlyAspAsnTrpGlyProleuValAlaValLysGlnleuGln 855
Db 2677 GCGGCTATGACCCGCTAGGCAACATACAGTGGCTGTGGCTGTGAACAGCTGCGACG 2736
Oy 855 lsservalProAspGlnGlnArgAspPheGlnArgGlyIleGlnIleleuLysAlaLeuH 875
Db 2737 ACACGCGGCGACAGACAGAGAGGACTTTCAGCGGAGATTCAGATCTCAAGAGCAGC 2796
Oy 875 lsservalPheIleValLysTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeu 895

```

```

Db 2797 ACATGATTCATATGTCAAGATATCTGTGTCTACSTATATGGCCGCGGCGACAGCCTGTG 2856
Oy 895 rgleuValMetGlnTyrleuProSerGlyCysLeuArgAspLeuGlnArgHlaSerG 915
Db 2857 GCGTGTATGAGTACCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2916
Oy 915 lyleuNlaSerAspArgleuLeuPheAlaTrpGlnIleCysLysGlyMetGlyTyrL 935
Db 2917 GCGTGTATGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2976
Oy 935 euGlyAlaArgGlyCysValHlaAspArgleuAlaAlaArgAsnIleleuValGluSerG 955
Db 2977 TGGGCTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3036
Oy 955 lualaHlaIleLysIleAlaAspPheGlyLeuAlaLysleuLeuProleuGlyLysAspT 975
Db 3037 AGGCACAGCTCAAGATCTGCTAGCTTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3096
Oy 975 yTyrValValArgGluProGlyGlnSerProIlePheTrpTyrAlaProGlnSerLeu 995
Db 3097 ACTAGCTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3156
Oy 995 eAspAsnIlePheSerArgGlnSerAspValTrpSerPheGlyValValleuTyrGln 1015
Db 3157 CCGACACATCTTCTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3216
Oy 1015 ePheThrTyrCysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetGlyP 1035
Db 3217 TCTTACCTACTAGCGACAAAGCTGACGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3276
Oy 1035 roGlnArgGlyGlyProProleuCysArgleuLeuGlnleuLeuAlaGlnArgArgGly 1055
Db 3277 GTGACCGGAGATGTCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3336
Oy 1055 euProProProProProProProProProProProProProProProProProPro 1065
Db 3337 TGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3368

```

RESULT 3

US-09-003-289-1

Sequence 1, Application US/09003289

Patent No. 5916792

GENERAL INFORMATION:

APPLICANT: Civin, Curt I.

APPLICANT: Small, Donald

TITLE OF INVENTION: NOVEL PROTEIN TYROSINE KINASE, JAK3

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson P.C.

STREET: 4225 Executive Square, Suite 1400

CITY: La Jolla

STATE: CA

COUNTRY: USA

ZIP: 92037

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/003, 289

FILING DATE:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/357, 598

FILING DATE: 15-DEC-1994

ATTORNEY/AGENT INFORMATION:

NAME: Haile, Lisa A.

REGISTRATION NUMBER: 38,347

REFERENCE/DOCKET NUMBER: 07265/033001

TELECOMMUNICATION INFORMATION:

TELEPHONE: 619/678-5070

TELEFAX: 619/678-5099

; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 3807 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA
 ; US-09-003-289-1

Alignment Scores:

Pred. No.:	0	Length:	3807
Score:	4418.00	Matches:	870
Best Local Similarity:	86.94%	Conservative:	62
Query Match:	81.16%	Mismatches:	130
	75.39%	Indels:	12
		Gaps:	6

US-09-397-967-16 (1-1099) x US-09-003-289-1 (1-3807)

```

QY 1 MetAlaProSerGluGluThrProLeuIleProGlnArgSerCysSerLeuSer 20
   |||||
Db 168 ATGGACCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 227
   |||||
QY 21 SerGluAlaGlyAlaLeuHisValLeuLeuProProArgGlyProGlyProGlnArg 40
   |||||
Db 228 ACGGAGGCTGGGCTCCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 287
   |||||
QY 41 LeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuValArgAlaAlaVal 60
   |||||
Db 288 CATCTCTTCTCTTGGGAGCACTTGGCTGAGAGAGAGAGAGAGAGAGAGAGAG 347
   |||||
QY 61 CysGlyIleLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSer 80
   |||||
Db 348 ACGGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 407
   |||||
QY 81 ThrPheProSerHisIlePheCysIleGluAspValAspThrGluValLeuVal 100
   |||||
Db 408 TGGTT-CCCCGAGCCACATCTTCCGAGAGATGCCAGCACCCCAATCCTGCTGTA 466
   |||||
QY 100 rArgLeuArgPheTyrPheProAspTrpPheGlyLeuGluThrCysHisArgPheGly 120
   |||||
Db 467 CAGGATTCCTTACTTCCCAATTGGTTGGGCTGGAGAAGTGGCCACCGCTTGGGCT 526
   |||||
QY 120 uArgGlyAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLeuPheAla 140
   |||||
Db 527 ACGCAAGGATTTGGCCAGTCTTCTGACCTGACGAGCTGCTGCTGCTGCTGCTG 586
   |||||
QY 140 nHisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLysGluGln 160
   |||||
Db 587 GCACCCCACTGACCTGTGAGTGGGCGCTCCCGCTGAGCTCAGTCTCAAGAGCA 646
   |||||
QY 160 yGluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGlnAla 180
   |||||
Db 647 TGAAGTCTCAGCGCTGGCGTGTGACCGCGCGGATGGCGGAGAGAGAGAGAG 706
   |||||
QY 180 gProGlyGluLeuLeuLysThrValSerTyrLysAlaCysLeuProProSerLeuArg 200
   |||||
Db 707 GCGGAGAGAGCTGCTAGAGCTGCTAGCTACAGAGCGCTGCTGCTGCTGCTGCTG 766
   |||||
QY 200 pValIleGlnGlyGlnAspPheValThrArgArgGlyLeuArgGlyThrValValLeu 220
   |||||
Db 767 CCGATGCCAGGCGCTGCTGCTGAGCGGAGCGGATTCGAGAGAGAGAGAGAG 826
   |||||
QY 220 aLeuLeuProCysGlyArgLeuProGlyArgPro-TyrAlaLeuMetAlaLysTyrIle 240
   |||||
Db 827 CCTGGC-CCGGGAGGCGCGCTGCCAGGAGAGAGAGAGAGAGAGAGAGAGAGATCA 885
   |||||
QY 240 euAspLeuGluArgLeuHisProAlaAlaThrThrGluThrPheArgValGlyLeuPro 260
   |||||
Db 886 TGGAGCTGGAGGCGGTGATCCAGCGGAGGCGCGGAGAGAGAGAGAGAGAGAG 945
   |||||
QY 260 yALA-----GlnGluGluProGlyLeuLeuArgValAlaGlyAspAsnGlyIle 277
   |||||

```

```

Db 946 GGGCCCTTGGTGGCCAGCAGGAGGCTGGGCTGCTCCGCTGCTGTGAGAGGCGCATCG 1005
   |||||
QY 277 rPhePheSerAspAsp---GluLeuPheGlnThrPheCysAspPheProGluIleVal 296
   |||||
Db 1006 CTTGGACCCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1065
   |||||
QY 296 sPValSerIleAsnGlnAlaProArgValGlyProAlaGlyGluHisArgLeuValThr 316
   |||||
Db 1066 ACATTACATCAAGCAGAGCCCGGCGCTGGCCCGGCGGAGAGAGAGAGAGAGAG 1125
   |||||
QY 316 aThrArgMetAspGlyHisIleLeuGluAlaGluPheProGlyLeuProGluAlaLeu 336
   |||||
Db 1126 TTAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1185
   |||||
QY 336 ePheValAlaLeuValAspGlyTyrPheArgLeuIleCysAspSerArgHisTyrPhe 356
   |||||
Db 1186 GCTTGGGCGCTGCTGAGGCTACTTCCGGCTGAGACAGAGAGAGAGAGAGAGAG 1245
   |||||
QY 356 yLysGluValAlaProProArgLeuGluGluGluAlaAspValCysHisGlyPro 376
   |||||
Db 1246 GCAAGGAGGTG---GACCCAGGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAG 1302
   |||||
QY 376 lThrIleAspPheAlaIleHisLysLeuLysAlaAlaGlySerLeuProGlyThrTyr 396
   |||||
Db 1303 TCACCTGGAAGCTTGGCCATCAAGCTCAAGAGAGAGAGAGAGAGAGAGAGAGAG 1362
   |||||
QY 396 lLeuArgArgSerProGlnAspTyrAspSerPheLeuLeuThrAlaCysValGlnThr 416
   |||||
Db 1363 TTCTCCCGCCCATCCCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1422
   |||||
QY 416 rLeuGlyProAspTyrLysGlyCysLeuIleArgGlnAspProSerGlyAlaPheSer 436
   |||||
Db 1423 CCTTGGCTGCTGATTAAGAGGCTGCTCATCCGCGCACCCAGAGAGAGAGAG 1482
   |||||
QY 436 euValGlyLeuSerGlnProHisArgSerLeuArgGluLeuAlaAlaCysTrpAsn 456
   |||||
Db 1483 TGGTGGCTCAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1542
   |||||
QY 456 eArgGlyLeuArgValAspGlyAlaAlaLeuThrLeuThrSerCysAlaProArgPro 476
   |||||
Db 1543 GGGGCTGACAGTATGATGGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1602
   |||||
QY 476 yGluLysSerAsnLeuIleValValArgArgGlyCysAsnProAlaProAlaPro 496
   |||||
Db 1603 AAGAAAGTCCAACTGATGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 1662
   |||||
QY 496 ySerPro---SerCysAlaAlaLeuThrGlnLeuSerPheHisThrIleProThrAsp 515
   |||||
Db 1663 TTACAGCCCATCCCATACAGCTGAGTCAAGATGATTCACAAAGATCCCTGCTACA 1722
   |||||
QY 515 eLeuGluThrPheHisGluAsnLeuGlyHisGlySerPheThrLysIlePheArgGlySer 535
   |||||
Db 1723 GCGTGGAGTGGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1782
   |||||
QY 535 rGArgGluValValAspGlyGluThrHisAspSerGluValLeuLeuLysValMetAsp 555
   |||||
Db 1783 GCATGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1842
   |||||
QY 555 eArgHisArgAsnCysMetGluSerPheLeuGluAlaIleSerLeuMetSerGlnVal 575
   |||||
Db 1843 CCAAGCACAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1902
   |||||
QY 575 eTyrProHisLeuValLeuLeuHisGlyValCysMetAlaGlyAspSerIleMetVal 595
   |||||
Db 1903 CGTACCGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1962
   |||||
QY 595 lGluPheValTyrLeuGlyAlaIleAspMetTyrLeuArgGlyAspGlyHisLeuVal 615
   |||||
Db 1963 AGGAATTTGATACACTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2022
   |||||
QY 615 eAlaSerTyrLysLeuGlnValThrLysGlnLeuAlaTyrAlaLeuAsnTyrLeuGlu 635
   |||||
Db 2023 CAGCCAGCTGGAAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2082
   |||||

```

```

Qy 635 sPLySGlyLeuPronHISglYAsnValSerAlaArgLySValLeuAlaArgLUGlyG 655
    |||||
Db 2083 ACAAGAGCCGTGTCATGGCAATGTCTGCCCCGGAAGGTGCTCGGCTGGAGAGGGG 2142
Qy 655 lYAspGLyAsnProPhePheIleYLeuSerAspProGLyValSerProthrValLeuS 675
    |||||
Db 2143 CTGATGGAGGCCGCTTCATCAACGTGATGACCTGGGCTGACGCCGCTGTGTA 2202
Qy 675 eRLeuGLuLeuLeuThrAspArgLLeProTPrValAlaProGLyCysLeuGLInGLuAlaG 695
    |||||
Db 2203 GCCTGGAGATGCTACCGCAGAGATCCCTGGGTGGCCCCCGAGGTCTCCGGAGAGCCG 2262
Qy 695 lntHLeuCysLeuGLuAlaAspLySTPrGLyPheGLyValThrThrTrGLuValPheG 715
    |||||
Db 2263 ACACACTTAgCTTGGAACTGACAAAGTGGGGCTGGGCCCAAGCTCGGAAGTGTTA 2322
Qy 715 lntArgGLyProAlaHISleIleThrSerLeuGLuProAlaLysLysLeuLysPheTYrGLuA 735
    |||||
Db 2323 GTGGCGTCACCATGCCCATCAGTGGCTGATGCTGCTAAGAACTCCAAATTTTATGAGG 2382
Qy 735 sPGLnGLyGLInLeuProAlaLeuLysTrPrThrGLuLeuAlaGLyLeuIleThrGLInCysM 755
    |||||
Db 2383 ACCGGAGCAGCTGTGGGCCCCCAAGTGGACAGAGCTGGCCCTGATTCACAGAGTGA 2442
Qy 755 eValATyAspProGLyArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGLyL 775
    |||||
Db 2443 TGGCCTATAGCCGGGTGACAGAGGCCCTCTTACGAGCCGTACATGGTACCTCAATAGTC 2502
Qy 775 euIleThrSerAspTYrGLuLeuLeuSerAspProThrProGLyIleProSerProArgA 795
    |||||
Db 2503 TCATCTCTTCAAGATAGTGGCTCTCTGACACACACC---TGGTGCCCTGGCACCTGGT 2559
Qy 795 sPGLyLeuCysValAlaGLyAlaGLInLeuTYrAlaCysGLInAspProAlaIlePheGLInG 815
    |||||
Db 2560 ATGGGGCTGG---AATGGTGCCAGCTCTATGCTGCCCAAGACCCACAGATCTTCGAGG 2616
Qy 815 lntArgHISLeuLysTYrIleSerLeuGLyLysGLyAsnPheGLySerValGLyLeuC 835
    |||||
Db 2617 AAGAGACCTCAAGATACATCTCACAGCTGGGCAAGGCCCTTCTTGGCACCGTGGAGCTGT 2676
Qy 835 yAsArgTYrAspProLeuGLyAspAsnThrGLyProLeuValAlaValLysGLInLeuGLInH 855
    |||||
Db 2677 GCGGCTATACCCCGCTAGCGGCAATACAGGTGGCTGGTGGCCCTGGAACAGCTGCAGC 2736
Qy 855 lAsSerValProAspGLInGLInArgAspPheGLInArgGLyIleGLInLeuLysAlaLeuH 875
    |||||
Db 2737 ACAGCGGGCCAGACAGGAGGAGGACTTTCAGCGGAGATTCAGATCTCAAGACACAGC 2796
Qy 875 lAsSerAspPheIleValLysTYrArgGLyValSerTYrGLyProGLyArgGLInSerLeuA 895
    |||||
Db 2797 ACAGTGAATTCATGTCTCAAGTATCTGTGTGAGCTATGGCCGGGCCCGCCAGACCCCTG 2856
Qy 895 rGLuValMeGLuTYrLeuProSerGLyCysLeuAlaArgAspLeuGLInArgHISArgG 915
    |||||
Db 2857 CGCTGTGTCATGAGTACCTGCCAGCGGTGCTGCGGACCTCTCTGACAGCGGACCGGG 2916
Qy 915 lYLeuHISThrAspArgLLeuLeuPheAlaTrPrGLInIleCysLysGLyMeGLuTYrL 935
    |||||
Db 2917 GCCTGATGTCAGCGGCTCTCTGTATCTCTCCAGATCTGCAAGGGCAGTGGATACC 2976
Qy 935 euGLyAlaArgArgCysValHISArgAspLeuAlaAlaArgAsnIleLeuValGLuSerG 955
    |||||
Db 2977 TGGGCTCCCGGCGCTGGCTGACCGGACCTGGCGCGCGGAAACATCTCTGGTGGAGAGCG 3036
Qy 955 lAlaHISValLysLysIleAlaAspPheGLyLeuAlaLysLeuLeuProLeuGLyLysAspT 975
    |||||
Db 3037 AGGACAGAGTCAAGATCCGTGCTGCGGCTAGCTAGCTGCTGCGGCTTGCACAAAGACT 3096
Qy 975 yTYrValValArgGLuProGLyGLInSerProIlePheThrPrTYrAlaProGLuSerLeuS 995
    |||||
Db 3097 ACTAGGTGTCGCGGAGGCGAGGCGGAGCCCATTTTCTGTGATGCCCGCGAATCCCTCT 3156

```

```

Qy 995 eRAspAsnIlePheSerArgGLInSerAspValTrPrSerPheGLyValLeuTYrGLuL 1015
    |||||
Db 3157 CGGACACACTTCTTCTGCCCCAGCTGACAGCTGTGGAGCTTCCGGGTCTGTCAGAGC 3216
Qy 1015 euPheThrTYrCysAspLysSerCysSerProSerAlaGLuPheLeuArgMeGLyP 1035
    |||||
Db 3217 TCTTCACCTACAGCAGCAAAAGCTGACGCCCTCGGCGGAGTTCCTGGCGGATGGAGT 3276
Qy 1035 roGLuArgGLuLysProProLeuCysArgLeuLeuGLuLeuAlaGLyAlaArgTYrGL 1055
    |||||
Db 3277 GTGACGGGAGTGTCCCGGCTCTGCGGCTCTTGTGAACCTGTGAGGAGGCGGAGAGGC 3336
Qy 1055 euProProProProthrCysProThrGLuVal 1065
    |||||
Db 3337 TGCCGGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3368

RESULT 4
PCT-US95-16435-1
; Sequence 1, Application PC/TUS9516435
; GENERAL INFORMATION:
; APPLICANT: The Johns Hopkins University School of Medicine
; TITLE OF INVENTION: NOVEL PROTEIN TYROSINE KINASE, JAK3
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/16435
; FILING DATE: 15-DEC-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Halle, Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07265/033WO1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3807 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; PCT-US95-16435-1

Alignment Scores:
Pred. No.: 0 Length: 3807
Score: 4418.00 Matches: 870
Percent Similarity: 86.94% Conservative: 62
Best Local Similarity: 81.16% Mismatches: 130
Query Match: 73.39% Indels: 12
DB: 5 Gaps: 6

US-09-397-967-16 (1-1099) x PCT-US95-16435-1 (1-3807)
Qy 1 MetaAspProSerGLuGLInThrProLeuIleProGLInArgSerCysSerLeuSerSer 20
    |||||
Db 168 ATGGACCTCCAAAGTGAAGAGAGCGCCCTGATCCCTCAGCGTTCAATGACAGCTCTTGCC 227
Qy 21 SerGLuAlaGLuAlaLeuHISValLeuLeuProProArgGLyProGLyProProGLInArg 40
    |||||
Db 228 ACAGAGGCTGTGCGCTCTCATGTGCTGCGGCTGGGGGCCCGGAGGCCCGCCCAAGCGC 287

```


Db 2443 TGGCCATATGACCGCGTCCAGAGCGCCCTCTTACAGACCGCTATTCGTGACCTCATATAGTC 2502
 QY 775 euilrhrseraspyrgluleuleuSeraprophrogllyleproserproara 795
 Db 2503 TCATCTTTCAGACTATGACCTCTCTCAGACACACAC--TGGTCCCTGGCAGCTCGTG 2559
 QY 795 spgluleuCyValAlaglyAlaglnleuYrAlaCysglInapProAlaIlePheglu 815
 Db 2560 ATGGCGCTGTGG--AATGGTGCCAGCTCTATGCTGCCAAGAGCCAGAGCTTTGGAGG 2616
 QY 815 luArghlaleuYsYrIleSerleuLeuGlyslYslnPheglYserValgluleu 835
 Db 2617 AGGAGACCCCTCAGTACATCTCAGCTGCGCAAGGGCTTCTTGGCAGCGTGAGACTGT 2676
 QY 835 ysArGTyrAspProleuGlyAspAsnThrGlyProleuValAlaValYsGlnleuGln 855
 Db 2677 GCCGCTATGACCGCGTACAGCAATACAGGTGCCCTGGTGGCCGCGGAACAGCTGCAGC 2736
 QY 855 lsserValProAspGlnAlaArgAspPheGlnArgGluIleGlnIleleuYsAlaLeu 875
 Db 2737 ACAGCGGGCCAGACAGCAGAGGAGCTTTCAGCGGAGATTCAGATCCCTCAAGACAGCAGC 2796
 QY 875 lsserAspPheIleValYsYrArgGlyValSerGlyProGlyArgGlnSerleu 895
 Db 2797 ACAGTATTCATTCATCAGATCTCGTGTGTACCTATGCGCGGCGCGCAGAGCCCTG 2856
 QY 895 rgluValMetGlyTyrleuProSerGlyCysleuArgAspLeuGlnleuGlnArg 915
 Db 2857 CCTGCTCATAGAGTACCTGCGCCAGCGGCTGCTGGCGACCTTCTGCGAGCGGACCGGG 2916
 QY 915 lYleuHlshThrsprArgleuLeuPheAlaIleTroglnleCyslyslYsGlyMetGlyTyr 935
 Db 2917 GCCTGATGCCAGCGCGCCCTCTCTATTCCTGCGAGATTCGCAAGGCGATGAGTAC 2976
 QY 935 euGlyAlaArgArgCysValAlaSerAspLeuAlaAlaArgAsnIleleuValIleuSer 955
 Db 2977 TGGGCTCCCGCGCTGCGTGACCGCGAGCTGCGCGCGGAGAACATCTCTGGAGAGCG 3036
 QY 955 luAlaHlshValYsIleAlaAspPheGlyleuAlaLysleuLeuProleuGlyLysAsp 975
 Db 3037 AGGCACAGCTCAAGATCGCTGCGCTTCCGCTAGCTGCTGCGCTTGCCTGCAAGAGCT 3096
 QY 975 YTYrValValArgIuPrgIuPrgIuPrgIuPrgIuPrgIuPrgIuPrgIuPrgIuPrgIuPrg 995
 Db 3097 ACTAGCTGGTCCGCGAGCGAGCGAGCCAGCCCATTTCTGTGATGCCCGCATCTCTCT 3156
 QY 995 eArAspAsnIlePheSerArgIuPrgIuPrgIuPrgIuPrgIuPrgIuPrgIuPrgIuPrg 1015
 Db 3157 CGGACAAATCTTCTCTGCGCGAGTCAAGCTGCGAGCTGCGGCGTCTGCTGAGAGC 3216
 QY 1015 euPheThrTyrCysAspLysSerCysSerProSerAlaGlnPheleuArgMetMetGly 1035
 Db 3217 TCTTACTACTGCGACAAAGAGCTGCGAGCTGCGCGAGTCTGCGGAGTATGAGGAT 3276
 QY 1035 roGluArgGlnIuPrgIuPrgIuPrgIuPrgIuPrgIuPrgIuPrgIuPrgIuPrgIuPrg 1055
 Db 3277 GTGACGGGAGATGCCCGCGCTGCGCGCTCTTGAACCTGTGAGAGAGGCGCAGAGC 3336
 QY 1055 euProProProProProProProProProProProProProProProProProProPro 1065
 Db 3337 TCGCGCGCGCTCTGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3368

RESULT 5

US-09-046-158A-21

Sequence 21, Application US/09046158A

Patent No. 6187552

GENERAL INFORMATION:

APPLICANT: Robert, Steven L.

APPLICANT: Kayles, Paul S.

TITLE OF INVENTION: METHOD FOR IDENTIFYING INHIBITORS OF

NUMBER OF SEQUENCES: 30

CORRESPONDENCE ADDRESS:

ADDRESS: Pharmacia & Upjohn Co., Intellectual Property
 ADDRESS: Legal Services
 STREET: 301 Henrietta Street
 CITY: Kalamazoo
 STATE: MI
 COUNTRY: USA
 ZIP: 49001
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/046,158A
 FILING DATE:
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Darlley Jr., James D.
 REGISTRATION NUMBER: 33,673
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 616/833-2210
 TELEFAX: 616/833-8897
 TELEX: 224401
 INFORMATION FOR SEQ ID NO: 21:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3435 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 US-09-046-158A-21
 Alignment Scores:
 Pred. No.: 7,7e-235 Length: 3435
 Score: 2628.00 Matches: 534
 Percent Similarity: 66.21% Conservative: 193
 Best Local Similarity: 48.63% Mismatches: 340
 Query Match: 44.86% Indels: 31
 Gaps: 14
 US-09-397-967-16 (1-1099) x US-09-046-158A-21 (1-3435)
 QY 20 SerGluAlaGlyAlaLeuHlshValleuLeuPro-ProArgGlyProGlyProProGly 39
 Db 111 TCCAGTCTTCAGGT-----GTATCTTACCATTCCCTGGAAATCTGAGCGAGA 161
 QY 39 nArgLeuSerPheSerPheGlyAspTyrleuAlaGluAspLeuValArgAlaAla 59
 Db 162 TTATCTGACCTTTCCATCTGGGAGTATGTGCAAGAAATCTGTATTCGCTCTTAA 221
 QY 59 sAlaCysGlyIleleuProValTyrHlshSerleuPheAlaLeuAlaThrGluAspPhe 79
 Db 222 AGCTTGTGTATCCACACTGTGTATCATATATGTTTGTATATGATGAGCAAG 281
 QY 79 rCysTyrPheProProSerHlshIlePheCysIleGluAspValAspThrGluVal 99
 Db 282 GATCTGATCCACCAACATGCTTCCATATGATGAGCAACAGGATATATGACT 341
 QY 99 lTYrArgLeuArgPheTyrPheProAspTyrPhe-----GlyleuGlnThrCysHlsh 117
 Db 402 GCATGGAATATCTCGAGGTGCTGAAGCTCCCTCTTGATGACTTTCATGCTTACT 461
 QY 117 gPheGlyleuArgLysAspLeuThrSerAlaIleleuAspLeuHlshValleuGlnHlsh 137
 Db 462 CTTCGCTCAGTGGCGGAGATGTTTGTGATGATGATGATGATGATGATGATGATGATGATGATG 521
 QY 157 sGluGlnGlyIuPheleuSerleuAlaValleuAspLeuAlaGlnMetAlaArgGlu 177
 Db 522 AACACAGAGAAATGCTTGTGGATGCGAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATG 581


```

|||||
Db 2730 CATTGTAAGTACAGGGAGTGTGCTACAGTCTGCTCGCGCTAACTAAATTAATATAT 2789
Oy 898 tGluTyrLeuProSerGlyCysLeuArgAspLeuGluArgHisArg---GlyLeuH 917
Db 2790 GGAATATTACCATATGAGAGTTTACGAGACTATCTTCAAAACATAAAGACGGATAA 2849
Oy 917 sThrAspArgLeuLeuLeuPheAlaTrpGlnIleCysGlyMetGluTyrLeuGlyAl 937
Db 2850 TCACATAAACCTTTCGACATACACATCTGACATATGCAGAGGATAGAGACTATCTTGTC 2909
Oy 937 sArgArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGluSerGluAlaH 957
Db 2910 AAAAAGGATATATCCACAGGAGATCTGGCAACGAGAAATATATTGGTGAGAACAGAGAAC 2969
Oy 957 sValIysIleAlaAspPheGlyLeuAlaIleAspLeuProLeuGlyLysAspTyrTyrVa 977
Db 2970 ACTTAAATTTGAGATTTTGGGTTAACCAAGCTTGGCCACAGCAAGCAAGAAATACTATTA 3029
Oy 977 lValArgGluProGlyGlnSerProIlePheTrpTyrAlaProGluSerLeuSerAspAs 997
Db 3030 ACTAAAGAACCTGGTGAAAGTCCCATATTTGTATGCTCCAGATCTGACAGAGAG 3089
Oy 997 nIlePheSerArgGlnSerAspValTrpSerPheGlyValIleLeuTyrGluLeuPheH 1017
Db 3090 CAGTTTCTGTGCGCTCAGATGTTTGAGCTTGGAGTGTCTGTATGAACCTTTTCAC 3149
Oy 1017 rTyrCysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetGlyProGluArg 1037
Db 3150 ATACATTGAGAGAGTAAAGTCCACACAGCGAATTTATGCTATGATTTGGCAATGACAA 3209
Oy 1037 gGluGlyProProLeuCys---ArgLeuLeuGluLeuLeuAlaGluGlyArgArgLeuPr 1056
Db 3210 ACAAGCACAGATAGTGTGCTTCCATTTGATGACACTTTTGAAACATTAATGAAGATTACC 3269
Oy 1056 oProProProThrCysProThrGluValGlnGluLeuMetGlnLeuCysTrpAlaProGl 1076
Db 3270 AAGACCAAGTGAATGCCCGATGAGATCTATATGATCATGACAGAGATGCTGGAAACAATA 3329
Oy 1076 uProHisAspArgProAlaPheAlaThrLeuSerProGlnLeuAspProLeu 1093
Db 3330 TGTAAATCAACGCCCTCTTAGGAGTCTAGCTCTTGAGAGTGAATCAATA 3381

```

RESULT 6

```

US-08-567-508C-1
: Sequence 1, Application US/08567508C
: Patent No. 5914393
: GENERAL INFORMATION:
: APPLICANT: Coleman, Roger
: TITLE OF INVENTION: A NOVEL HUMAN JAK2 KINASE
: NUMBER OF SEQUENCES: 5
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
: STREET: 3174 Porter Drive
: CITY: Palo Alto
: STATE: CA
: COUNTRY: US
: ZIP: 94304
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FastSeq Version 1.5
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/567,508C
: FILING DATE: 05-DEC-1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Billings, Lucy J.

```

```

: REGISTRATION NUMBER: 36,749
: REFERENCE/DOCKET NUMBER: PF-00490S
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 650-855-0555
: TELEFAX: 650-845-4166
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 4482 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: IMMEDIATE SOURCE:
: LIBRARY: Placentia
: CLONE: 179527
: US-08-567-508C-1
Alignment Scores:
Pred. No.: 6,59e-234 Length: 4482
Score: 2621.00 Matches: 532
Percent Similarity: 66.03% Conservative: 193
Best Local Similarity: 48.45% Mismatches: 342
Query Match: 44.73% Indels: 31
Db: 2 Gaps: 14
US-09-397-967-16 (1-1099) x US-08-567-508C-1 (1-4482)
Oy 20 serserGluAlaGlyAlaLeuHisValLeuLeuPro-ProArgGlyProGlyProProGl 39
Db 447 TTCAGTTCTTCAGGT-----GTATCTTTACCATTTCCCTTGGAAATCTGAGCGACA 497
Oy 39 nArgLeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysValArgAlaAla 59
Db 498 TTATCTGACCTTTTCATCTGCGGAGATGTTGAGAGAAATCTGATTCGCTTCATA 557
Oy 59 sAlaCysGlyIleLeuProValTyrHisSerIleuPheAlaLeuAlaThrGluAspPhe 79
Db 558 ACCTTGTGTATCACACCTGTGTATCATAAATATGTTGCTTAAATGAGTGAAGAAAG 617
Oy 79 rCysTrpPheProProSerHisIlePheCysIleGluAspValAspThrGlnValLeuVa 99
Db 618 GATCTGTATCCACCACCAACCATGCTTCTCATATAGATGATGACAGGCAATATGACT 677
Oy 99 lTyrArgLeuArgPheTyrPheProAspTrpPhe-----GlyLeuGluThrCysHisAr 117
Db 678 CTACAGAAATTAAGATTCTTCTTCGTTGATGATGATGATGATGATGATGATGATGATG 737
Oy 117 gPheGlyLeuArgGlyAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLe 137
Db 738 GCATGGAATATCTGAGGTGCTGAAGCTCCCTCTTGTGATGACTTTCATGCTCTTACCT 797
Oy 137 uPheAlaGlnHisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLy 157
Db 798 CTTTGTCTGATGCGCGCATGATTTTGTGCATGATGATGATGATGATGATGATGATGAT 857
Oy 157 sGluGlnGlyIlePheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGlu 177
Db 858 AACACAGAGAAATGATCTGGATGATGATGATGATGATGATGATGATGATGATGATGATG 917
Oy 177 nAlaGlnArgProGlyGluLeuLeuLysThrValSerTyrLysAlaCysLeuProPse 197
Db 918 CGATCAAAACCCACTGCGCATCTAATACCTCTTCAGCTCAAGACATCTTCCACAAATG 977
Oy 197 rLeuArgAspValIleGlnGlyGlnAspPheValHisArgArgArgArgArgArgTrpVa 217
Db 978 TATTTGAGCAAGATCCAGAGATCATATTTTGTGACAGAGAGCAATTAAGTACAGATT 1037
Oy 217 lValLeuAlaLeuLeuProCysGlyArgLeuProGlyArgProTyrAlaLeuMetAla 237
Db 1038 TCGCAGATTATTATTCAGCATTCAGCCAAATGCAGCAATGCAGCAATGCAGCAATGC 1097
Oy 237 sTyrIleLeuAspLeuGluArgLeuHisProAlaAlaThrThrGluThrPheArgValGl 257

```

Db 1098 GTATTCTATAAATCTGGAACCTGCGAGTCTGCTTCTACACAGAGAATTTGAGTAA 1157
 Oy 257 YLEUPROGLYALA-----GInGUlGUlUPROGLYLeuLeuArgYAlAlaGlyAs 273
 Db 1158 AGAAGCTGGAAGTGCCTTCAGAGTGAAGATTTTGGCAACCATATTAATTAAGTGA 1217
 Oy 273 pAANGlyIleProTTrSerSer-----AsnAspGluLe 284
 Db 1218 CGGTGGATTTCAGTGTCAAGAGGAAACATTAAGAAAGTGAACACTGACAGACAGCA 1277
 Oy 284 uPheGInThrPheCysAspPheProGUlIleValAspValSerIleAsnGUlAlaProAr 304
 Db 1278 TTATAGATTATATTCGATTTTCTCTATATATATGATGATAGATTAAGACAGCAACCA 1337
 Oy 304 gValGlyProAlaGlyGUlHisArgLeuValIthValThArgMetAspGlyHisIleLe 324
 Db 1338 AGAGGCT--TCAAATGMAAGCGAGTGTACATATCCATAGCAACATGTAATAAAATCT 1394
 Oy 324 uGUlAlaGUlPheProGUlLeuProGUlAlaLeuSerPheValAlaLeuValAspGlyTy 344
 Db 1395 GGAATATGAACCTTAGCTCATTAAGGAGAGCTTGTCTTGTGTCATTAATGATGATA 1454
 Oy 344 rPheArgLeuIleCysAspSerArgHisTyPheCysGlyGUlValAlaProProArgLe 364
 Db 1455 TTATAGATTATATTCGATTTTCTCTATATATGATGATAGATTAAGACAGCAACCA 1514
 Oy 364 uLeuGUlGUlGUlAlaAspAlaCysHisGlyProIleThLeuAspPheAlaIleHisLy 384
 Db 1515 GCTTGAATAATATACAAAGCAACTGTCATGGCCCAATTCGATGATTTGCTACTTAA 1574
 Oy 384 sLeuValAlaAlaGlySerLeuProGUlThrTyIleLeuArgArgSerProGUlAspTy 404
 Db 1575 ACAGAAACAAAGCAAGTATACAGACTGAGCTGATGATCTTGCATGAGCTCTTAAGCACT 1634
 Oy 404 rAspSerPheLeuLeuThralaCysValGlnThrProLeuGlyProAspTyIleGlyCy 424
 Db 1635 TAATTAATATTTTGTGCTTTGCTGTCGAGCAAGAAATGTCAATTAATAAATCACTG 1694
 Oy 424 sLeuIleArgGUlAspProSerGUlAlaPheSerLeuValGlyLeuSerGlnProHisAr 444
 Db 1695 TTGATATACAAAATAGATAGATGAAGTACAACTCAGTGGGCAAGAAAGAACTTCAG 1754
 Oy 444 gSerLeuArgGUlLeuLeuAlaAlaCysTrpAsnSerGlyLeuArgValAspGlyAlaAl 464
 Db 1755 CAGCTTAAAGATCTTTGATTTGATACAGATGAAGAACTGCTGCACAGCAATATATAT 1814
 Oy 464 aLeuTyIleThrSerGlyCysAlaProArgProGlyGUlGlySerAsnLeuIleVala 484
 Db 1815 TTTCAGATTACTAAATGCTGTCGCCCAAGCCAAAGATTAATCAAACTTCTAGTCTT 1874
 Oy 484 lArg---ArgGUlCysAsnProAlaProAlaProGUlCysSerProSerGlyCysAlaLe 503
 Db 1875 CAGAGCAATAGTGTCTTGTGATGATACCACTCACAACATTAAGAGGCTTACTATAT 1934
 Oy 503 uThrGlnLeuSerPheHisThrIleProThrAspSerLeuGUlTrpHisGUlAsnLeuG 523
 Db 1935 GAACCAATAGTGTCTTCACAAAATCAGAAATGAAGATTGTATTAATGAAGAACTTGG 1994
 Oy 523 yHisGlySerPheThrLySIIlePheArgGlySerArgArgGUlValValaAsp---GlyG 542
 Db 1995 CCAAGGCACTTTACAAAGATTTTAAAGCGGTACAGAGAGAGAGATAGAGAGCTGCTCA 2054
 Oy 542 uThrHisAspSerGUlValLeuLeuValMetAspSerArgHisArgAsnCysMetG 562
 Db 2055 ACTGCATGAACACAGATCTTTTAAAGTCTGTGATTAACACACAGAGACTATTCAGA 2114
 Oy 562 uSerPheLeuGUlAlaAlaSerLeuMetSerGUlValSerTyProHisLeuValLeuLe 582
 Db 2115 GCTTTCTTTTGAACAGACAGATATGATGACAGACTTCTCACAAGCATTTGTTTAA 2174
 Oy 582 uHisGlyValaCysMetAlaGlyAsp---SerIleMetValGlnGUlPheValTyIleG 601
 Db 2175 TTATGAGATATGTCTGTGTGAGACAGAAATATTTCTGTTGAGAGTTTGAATAATTTGG 2234

Oy 601 yAlaIleAspMetLeuThrLeuArgLySArgGlyHisLeuValSerAlaSerTrpLyLeuG 621
 Db 2235 ATCATGATATATATCTGTGAAGAAATTAATAATGTATATATATATATATGCAACTTCA 2294
 Oy 621 nValThrLySGLnLeuAlaTrpAlaLeuAsnTyIleLeuAspLySGLyLeuProHisG 641
 Db 2295 AGTTGCTTAACAGTGTGGCAGTGGCCATGATTTTCTAGAGAAACACCTTATTCATGG 2354
 Oy 641 yAsnValSerAlaArgLySValLeuLeuAlaArgGUlGly-----AspGlyAsnPr 659
 Db 2355 GAATGTATGTGCCAAATAATTTCTGTATACAGAGAAAGACAGAGACAGAAATCC 2414
 Oy 659 oProPheIleLyLeuSerAspProGUlValSerProThrValLeuSerLeuGluMetLe 679
 Db 2415 TCCCTTATATCAACTTATAGTATCTGTCATTAATGAGAAATCGTAATAATTTAAATTT 2534
 Oy 679 uThrAspArgIleProTrpValAlaProGUlCysLeuGlnGUlAlaGlnThrLeuCysLe 699
 Db 2475 TCAGAGAGAAATACATGGGTACACCTGATGATGAAATCGTAATAATTTAAATTT 2534
 Oy 699 uGUlAlaAspLySTrpGlyPheGlyAlaThrThrTrpGUlValPheGlnArgGUlProAl 719
 Db 2535 GGCACAGCAAAATGAGATTGTGTGACACTTGTGGAAATCGCAGTGGAGAGATTA 2594
 Oy 719 aHisIleThrSerLeuGUlProAlaLySLeuLySLeuPheTyGlyAspGlnGlyLe 739
 Db 2595 ACCCTTAGTGTCTGTGATTTCTCAAGAAAGCTAACATTTTATGAGATGGCATCGCT 2654
 Oy 739 uProAlaLeuLySTrpThrGlnLeuAlaGlyLeuIleThrGlnCysMetAlaTyAspTr 759
 Db 2655 TCCGTGCACCAAGAGGGGCAAAATTAGCAACCTTAATTAATGTATGATTATGAAAC 2714
 Oy 759 oGlyArgArgProSerPheArgValIleLeuArgAspLeuAsnGlyLeuIleThrSerAs 779
 Db 2715 AGATTTCAGGCTTCTTTCAGAGCATCATCAGATCTTAACAGTTGTTTACTCCAGA 2774
 Oy 779 pTyGluLeuLeuSerAspProThrProGUlIleProSerProArg---AspGluLeuCy 798
 Db 2775 TTATGACACTTTTAACAGA---AATGACATGTTACCAATATGAGATGATGCTGCTGGG 2831
 Oy 798 sValAlaGlyAlaGlnLeuTyIleAlaCysGlnAspProAlaIlePheGUlGlnArgHisLe 818
 Db 2832 GTTTTGTGGTCC---TTGAAAGCCGGGATCTTACACAGTTGTGAAGAGACATTT 2885
 Oy 818 uLySTrIleSerLeuLeuGlyLySGLyAsnPheGlySerValGlnLeuCysArgTyAs 838
 Db 2886 GAATTTCTACAGCAACTTGGCAAGGTATTTTGGAGTGTGAGATGTGCGGTATGA 2945
 Oy 838 pProLeuGlyAspAsnThrGlyProLeuValAlaValLySGLnLeuGlnHisSerValPr 858
 Db 2946 CCCCTACAGACACAMCACTGGGAGTGTGCGGTAAAGAGCTTCAGATATGACTGA 3005
 Oy 858 oAspGlnGlnArgAspPheGlnArgGUlIleGlnIleLeuLySAlaLeuHisSerAspPr 878
 Db 3006 AGAGCACTTAAGAGACTTTGAAGAGAAATTTGAATCTCTGAATCCCTACAGCATGCAA 3065
 Oy 878 eIleValLySTrArgLySGLyValSerTyGlyProGUlArgGlnSerLeuArgLeuValMe 898
 Db 3066 CATTTGTAAGTACAGAGGAGTGTCTACAGTGTGCTGCGGTATCTTAATAATTAATAT 3125
 Oy 898 tGluTyIleuProSerGlyCysLeuAlaArgAspLeuGlnArgHisArg---GlyLeuH 917
 Db 3126 GGAATATTTTACCAATATGAGATTACAGACTATCTTCAAAACATTAAGAACAGATAGA 3185
 Oy 917 sThrAspArgLeuLeuLeuPheAlaTrpGlnIleCysLySGLyMetGlnTyIleGlyAl 937
 Db 3186 TCACATTAACATTTTGGAGTACATCTCAGATATGCAAGGATGAGATGATCTTGGTAC 3245
 Oy 937 aArgArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGlnSerGUlAlaH 957
 Db 3246 AAAAAGATATTCACAGAGATCTGCGACAGAGAAATATATATGTTGAGAAACAGAACAG 3305


```

Db 1338 AGAGGCT---TCAAATGAACCCGAGTTGATCATATCCATAGACAGAGTGGTAAATCT 1394
Oy 324 uGIuAaGIuPheProGIuLeuProGIuAaLeuSerPheValAlaLeuValAspGIuTY 344
Db 1395 GGAAATTGAACCTAGCTCATTAAGGAAGCTTTGCTTCTGTCATTAATTAATGATGATA 1454
Oy 344 rPheArGIuLeuIleCysAspSerArgHisTYrPheCysIysGIuValAlaIleProProlGle 364
Db 1455 TTATGATTTAACTGCAGATGACATCATCTACCTCTGTAAGAGTAAAGACCTCCAGCCGT 1514
Oy 364 uLeuGIuGIuAaIleAspValCysHisGIuProIleThrIleuAspPheAlaIleHisIly 384
Db 1515 GCTTAAATAATATACAAACCACTGTCATGCCCCAATTTGATGATTTGGCATTAGTAA 1574
Oy 384 sLeuIysAlaIaGIuSerLeuProGIuTYrIleuIleuArgArGIuSerProGIuAspTY 404
Db 1575 ACTGAAGAAGACGATTAACAGACTGACCTGATGATGATGATGATGATGATGATGATGAT 1634
Oy 404 rAspSerPheLeuLeuThrAlaCysValGIuThrIleuGIuProAspTYrIysGIuTY 424
Db 1635 TAATAAATATTTTGGACTTTGCTGTCGAGCGAGAAATGTCATTGATTAATTAACACTG 1694
Oy 424 sLeuIleArgGIuAspProSerGIuAlaPheSerLeuValGIuLeuSerGIuProHisAr 444
Db 1695 TTTGATTAACAAAATGAAGATGAAGATGACAACTCAGTGGACAAAGAAACCTGAG 1754
Oy 444 gSerLeuArgGIuLeuLeuAlaIaCysTYrAsnSerGIuIysSerAsnLeuIleVala 464
Db 1755 CAGTCTTAAGATCTTTGAAATGTTTACCAATGGAAGAACTGCTCCACCAATTAAT 1814
Oy 464 aLeuTYrLeuThrSerCysCysAlaProArgProIysGIuIysSerAsnLeuIleVala 484
Db 1815 TTTCCAGTTTACTAATGCTGTCGCCCAAGCCAAAGATTAATCAACCTCTAGTCT 1874
Oy 484 lArg---ArgGIuCysAsnProAlaProAlaProGIuCysSerProSerCysAlaIle 503
Db 1875 CAGAAGCAATGGTGTTCATGATGATCCAACTCACCACCATTAACAGAGCCCTACTCATAT 1934
Oy 503 uThrIleuSerPheHisIleThrIleProThrAspSerLeuGIuThrPheHisIleAsnLeu 523
Db 1935 GAACCAATGGGTTCACAAATCAAGAAATGAAGATTTGATTAATGAAGACCTTGG 1994
Oy 523 yHisGIuSerPheThrIlePheArgGIuSerArgGIuValValAsp---GIuGI 542
Db 1995 CCAAGGACCTTTTACAAAGATTTTAAAGCGTACGACAGAGTGGAGTACAGCTGA 2054
Oy 542 uThrHisAspSerGIuValLeuLeuIysValMetAspSerArgHisArgAsCysMetGI 562
Db 2055 ACTGATGAACAGAGAGTCTTTAAAGTCTGGATTAAGCACACACAGAGACTATTCA 2114
Oy 562 uSerPheLeuGIuAlaIleSerLeuMetSerGIuValSerTYrProHisIleValLeu 582
Db 2115 GTTCTTTCTTGAAGCAGCAAGTATGATGACAGCTTTCACACACATTTGGTTTAA 2174
Oy 582 uHisGIuValCysMetAlaGIuAsp---SerIleMetValGIuGIuPheValTYrLeuGI 601
Db 2175 TTATGAGATGTCTGCTGTCGACAGAGATATTTGCTGTCAGAGCTTGTAAATTTG 2234
Oy 601 yAlaIleAspMetTYrLeuArgIysArgGIuHisLeuValSerAlaSerTYrIysLeuGI 621
Db 2235 ATACTAGATACATATCTGAAAGAAATTAATTAATTAATTAATTAATTAATTAATTA 2294
Oy 621 nValThrIysGIuLeuAlaIleValAlaLeuAsnTYrLeuGIuAspIysGIuLeuProHisGI 641
Db 2295 AGTTGCTAAACAGTTGGCATGGCCATGCTTTCTTAAGAAAGAAACACCTTATTCATG 2354
Oy 641 yAsnValSerAlaArgIysValLeuLeuAlaArgGIuGIuIy---AspGIuAsnPr 659
Db 2355 GAATGATATGCGCAAAATATCTGCTTATCAGAGAAAGACAGACAGAGAAATCC 2414
Oy 659 oProPheIleIysLeuSerAspProGIuValSerProThrValLeuSerLeuGIuMetIle 679

```

```

Db 2415 TCCTTCATCAAACTTAGTGCATCTGGCATTAATTAATTAATTAATTAATTAATTAAT 2474
Oy 679 uThrAspArgIleProTYrValAlaIleProGIuCysLeuGIuAlaGIuThrIleCysIle 699
Db 2475 TCAGAGAGAGATACAGTGGGATCCCTGAATGATGAATTAATTAATTAATTAATTAAT 2534
Oy 699 uGIuAlaIleAspIysTYrPheGIuAlaThrIleThrIleuValPheGIuArgGIuProAl 719
Db 2535 GGCACACACAAATGAGATTTGGTACCCTTTGTGGAAATGTCAGTGGAGAGATTA 2594
Oy 719 aHisIleThrSerLeuGIuProAlaIleIleuArgAspLeuAsnGIuLeuIleThrSer 739
Db 2595 ACCCTAAGTCTGCTGATCTCAAGAAAGCTTAATTAATTAATTAATTAATTAATTAAT 2654
Oy 739 uProAlaLeuIysTYrThrIleuValAlaGIuLeuIleThrIleCysMetAlaTYrAspPr 759
Db 2655 TCCTGCACCAAGTGGCAGATTAAGCAACCTTAATTAATTAATTAATTAATTAATTAAT 2714
Oy 759 oGIuArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGIuLeuIleThrSer 779
Db 2715 AGATTTCAGGCTTCTTCAAGACCATCATACGATCTTAACAGTTGTTTACTCCAGA 2774
Oy 779 pTYrGIuLeuLeuSerAspProThrProGIuIleProSerProArg---AspGIuLeuCY 798
Db 2775 TTATGAATTAATTAACAGAA---AATGACATGTACCAATTAATTAATTAATTAATTAAT 818
Oy 798 sValAlaGIuAlaGIuLeuTYrAlaCysGIuAspProAlaIlePheGIuGIuArgHisIle 838
Db 2832 GTTTCTGCTGCTCC-----TTTGAAGACCGGGAGCTTCAACAGTTTGAAGACATTT 2885
Oy 818 uIysTYrIleSerLeuLeuGIuIysGIuAsnPheGIuSerValGIuLeuCysArgTYrAs 838
Db 2886 GAATTTTTCAGCAACTGGGCAAGGATTAATTTGGAGTGTGAGATGTCCCGTATGA 2945
Oy 838 pProLeuGIuAspAsnThrGIuProLeuValAlaValIysGIuLeuGIuHisSerValPr 858
Db 2946 CCTCTACAGACAAACATGGGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3005
Oy 858 oAspGIuIleArgAspPheGIuIleArgGIuIleGIuIleLeuIysAlaLeuHisSerAspPr 878
Db 3006 AGAGCACTTAAGAGCTTTGAAAGGAAATTAATTAATTAATTAATTAATTAATTAATTA 3065
Oy 878 eIleValIysTYrArgGIuValSerTYrGIuProGIuArgIleSerLeuArgLeuValMe 898
Db 3066 CATTTGAAGTACAGAGGAGTGTCTACAGTGTGTGTGGGTATTAATTAATTAATTAAT 3125
Oy 898 lGIuTYrLeuProSerGIuCysLeuArgAspLeuLeuGIuArgHisArg---GIuLeuH 917
Db 3126 GGAATATTATACCATATGGAAGTTTACGAGACTATCTCAAAAACATTAAGAAAGGATGA 3185
Oy 917 sThrAspArgLeuLeuLeuPheAlaIleTrpIleIleCysIysGIuMetGIuTYrLeuGIuAl 937
Db 3186 TCACATTAATACTTGCAGCTACATCTCAGATATGCAAGGATGAGATCTTGGTAC 3245
Oy 937 aArgArgCysValHisArgAspLeuAlaIleArgAsnIleLeuValGIuSerGIuAlaIleH 957
Db 3246 AAAAGGTATATCCACAGGATCTGGCAAGCAAGAAATATATTGTGGAGAACAGAACAG 3305
Oy 957 sValIysIleAlaAspPheGIuLeuAlaIysLeuLeuProLeuGIuIysAspTYrTYrVa 977
Db 3306 AGTTAAATTTGGRATTTTGGCTTAACCAAACTTTCGCCACAAAGCAAGAAATTAATTA 3365
Oy 977 lValArgGIuProGIuGIuSerProIlePheThrTYrAlaIleProGIuSerLeuSerAsp 997
Db 3366 AGTAAAGAACTGGGAAAGTCCCATATTTGATGTGCTCAGAAATCACTACAGAGAGAG 3425
Oy 997 nIlePheSerArgGIuSerAspValTrpSerPheGIuValValLeuTYrGIuLeuPheTh 1017
Db 3426 CAAGTTTCTGTGGCTCAGATGTTGGAGCTTTGGAGAGTGTCTGTAAGACTTTTCA 3485
Oy 1017 rTYrCysAspIysSerCysSerProSerAlaGIuPheLeuArgMetMetGIuProGIuAr 1037
Db 3486 ATACATTTAGAGAGATTAAGATCCACAGCGGAATTTATGCTATGATTTGCAATGACAA 3545

```

QY 1037 gclugluprproleucys---ArgleuLeuLeuLeuAlaGluGlyArgArgLeuPr 1056
 Db 3546 ACAAGACAGATGATGCTGTCATTTGATAGACCTTTTGAAGATTAAGAAAGATTACC 3605
 QY 1056 orprorprorprCysprrothGluValGlnLeuMetGlnLeuCysTrpAlaProG1 1076
 Db 3606 AAGACCGATGGATGCCGATGATGATCTATGATCATGACGAAATGCTGGAACAATTA 3665
 QY 1076 utronhAspArgProAlaPheAlaThrLeuSerProGlnLeuAspProLeu 1093
 Db 3666 TGTAAATCAAGCCCTCTTAAAGGATCTAGCTTTCGATGATCAATA 3717

RESULT 8
 US-08-097-997A-8
 : Sequence 8, Application US/08097997A
 : Patent No. 5728536
 : GENERAL INFORMATION:
 : APPLICANT: Inle, James N.
 : APPLICANT: Silvenoinen, Olli
 : APPLICANT: Mituhh, Bruce A.
 : APPLICANT: Ouelle, Frederick W.
 : TITLE OF INVENTION: Jak kinases and Regulation of Cytokine Signal
 : NUMBER OF SEQUENCES: 14
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Sterne, Kessler, Goldstein & Fox
 : STREET: 1100 New York Avenue, Suite 600
 : CITY: Washington
 : STATE: D.C.
 : COUNTRY: U.S.A.
 : ZIP: 20005-3934
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: Patentin Release #1.0, Version #1.25
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/097, 997A
 : FILING DATE: 29-JULY-1994
 : CLASSIFICATION: 435
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Fox, Samuel L.
 : REGISTRATION NUMBER: 30,353
 : REFERENCE/DOCKET NUMBER: 0656, 0370000/SLE/GKT
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: (202) 371-2600
 : TELEFAX: (202) 371-2540
 : INFORMATION FOR SEQ ID NO. 8:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 3629 base pairs
 : TYPE: nucleic acid
 : STRANDEDNESS: single
 : TOPOLOGY: linear
 : FEATURE:
 : NAME/KEY: CDS
 : LOCATION: 94...3480
 : US-08-097-997A-8

Alignment Scores:
 Pred. No.: 6.47e-234 Length: 3629
 Score: 2619.50 Matches: 537
 Percent Similarity: 65.50% Conservative: 190
 Best Local Similarity: 48.38% Mismatches: 350
 Query Match: 44.70% Indels: 33
 DB: 1 Gaps: 15

US-09-397-967-16 (1-1099) x US-08-097-997A-8 (1-3629)

QY 4 ProSerGluGluThrProLeuLeuProGlnAlaSerCysSerLeuSerSerGluAla 23
 Db 145 CCTGTACATCAAGATGGTATATTCCTGCAAGTCTTAATTCGTGAGAGATAGAGCA 204

QY 24 GluAlaLeuHisValLeuLeuProProArgGlyProGlyProProGlnArgLeuSerPhe 43
 Db 205 GTGCTTCAAGTATCTGATGACCATCTCTGTTGGCAAGCTGAAGAGATATGTGAAGTTT 264
 QY 44 SerPheGlyAspTrpLeuAlaGluAspLeuCysValArgAlaAlaLysAlaCysGlyLe 63
 Db 265 CCAAGTGGAGAGATATGTTGCACAGAAATTTGTGGGGCTGCTTGAAGCTTGATTT 324
 QY 64 LeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSerCysTrpPhePro 83
 Db 325 ACGCTGTGTATCATTAATATGTTTCCGTTTAATGAGTGAACCGAAAGCATGCTGATCCCA 384
 QY 84 ProSerHisIlePheCysIleGluAspValAspTrpGlnValLeuValTyrArgLeuArg 103
 Db 385 CCCATTCATGCTCTTGCACATAGACGATCAACAGCATGACATCTGTACAGGATTAAG 444
 QY 104 PheTyrPheProAspTrpPhe-----GlyLeuGluThrCysHisArgPheGlyLeuArg 121
 Db 445 TTCTACTTCCCTCATTTGAGTGTAGTGTAGTGACACAGACAGAACTACAGATACGAGTGC 504
 QY 122 LysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGlnHisLeuPheAlaGlnHis 141
 Db 505 CTTGGGGCTGAAGCTCTCTGCTTGAAGCATTTTGTATGCTTTTACCTTTTTCCTCAGTGC 564
 QY 142 ArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLysGluGlnGlyGlu 161
 Db 565 CGGCATGATTTTGTTCACAGGATGGATTAAGAACTGTGATCATGAATCAAGAAAG 624
 QY 162 PheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGlnAlaGlnArgPro 181
 Db 625 TGTCTTGGAGTGGCGGTGTAGACATGATGAATGAATGAAGAAAGAAAGACAGCTCCA 684
 QY 182 GluGluLeuLeuLysThrValSerTyrLysAlaCysLeuProProSerLeuArgAspVal 201
 Db 685 CTGGCTGTATTAACCTGTGCTGCTACACAGATCTTCAAGATGGCTGACCGAAG 744
 QY 202 IleGlnGlyGlnAspPheValThrArgArgArgIleArgArgThrValValLeuValLeu 221
 Db 745 ATCCAAAGATATACATTTTAACCCGGAAGCAATCAGATACAGATTTCCGAGATTCATT 804
 QY 222 LeuProCysGlyArgLeuProGlyArgProGlyArgValAlaLeuLysTyrIleLeuAsp 241
 Db 805 CAGCAATTCAGTCAATGTAAAGCCAGCTCCAGAAACCTTAACCTTAAGTATCTTAAAC 864
 QY 242 LeuGluArgLeuHisProAlaAlaThrThrGluThrPheAspVal----- 256
 Db 865 CTGGAAGACCTGACATGCTGCTTACACAGAAAGCTTTGAAGTAAAGAAATCTGCACAA 924
 QY 257 GlyLeuProGlyAlaGlnGluGluProGlyLeuLeuArgValAlaGlyAspAsnGlyLe 276
 Db 925 GGT---CCTTCAGTGGAGAGATTTTTCGAACCATTAATTAATACAGAAAGCGTGAAATT 981
 QY 277 ProTrpSerSer-----AsnAspGluLeuPheGlnThr 287
 Db 982 CAGTGGTCAAGAGGAACATTAAGGAAGTGAAGACCTGACAGACAGACAGCTACAGTTA 1041
 QY 288 PheCysAspPheProGluIleValAspValSerIleAsnGlnAlaProArgValIlePro 307
 Db 1042 TATTGTGATTTCCCTGATATTTATTTGATGTCACTATTAAAGCAAGCAACAGC---GAATGC 1098
 QY 308 AlaGlyGlnHisArgLeuValThrValThrArgMetAspGlyHisIleLeuGlnAlaGlu 327
 Db 1099 TCAATGAAGATGAAGATTTGATCACTGCATTAACAAGATCGTAAGATTTTGGAGATAGA 1158
 QY 328 PheProGlyLeuProGlnAlaLeuSerPheValAlaLeuValAspGlyTyrPheArgLeu 347
 Db 1159 CTTAGCTCATTAAGAAGCCCTTGATTCGTTCAATTAATTAATGAAGGGGTATTAACAGACTA 1218
 QY 348 IleCysAspSerArgHisTyrPheCysLysGluValAlaProProArgLeuLeuGluGlu 367
 Db 1219 ACTGGGATGGCCACCATTTACCTGTGCAAAAGAGTGGCTCCCGCCAGCTGTGCTGCAAGAC 1278
 QY 368 GluAlaAspValCysHisGlyProIleThrLeuAspPheAlaIleHisLysLeuValAla 387

Db 1279 ATACACAGCAACTGCCACGGCCCAATATCATGATTTTGGCATTTAGCAAACTAAGAG 1338
OY 388 AlAGlySerLeuProGlyThrTyrTlleuArgArgSerProGlnAspTyrAspSerPhe 407
Db 1339 GCGGGTAACGAGCTGATATGCTGTACAGTACGCCCTTAAGAGCTTCAACAAATAC 1398
OY 408 LeuLeuThrAlaCysValGlnThrProLeuGlyProAspTyrLysGlyCysLeuIleArg 427
Db 1399 TTTTGTACCTTGTCTGTGACCGGAAAAATGTCATTTGATTTAAACACTGTTTATTAACG 1458
OY 428 GlnAspProSerGlyAlaPheSerLeuValGlyLeuSerGlnProHisArgSerLeuArg 447
Db 1459 AAGAAATGAGATGAGATATACACCTGACGGGACTAAGAGAACTTCAGTAACCTTAAG 1518
OY 448 GluLeuLeuAlaAlaCysTyrPasnSerGlyLeuArgValAspGlyAlaIleLeuTyrLeu 467
Db 1519 GACCTTTTGAATTCCTACACGATGAAACGTGGCTCAGACATATCATCTTCACGTTT 1578
OY 468 ThrSerCysAlaProArgProLysGlyLysSerAsnLeuIleValAlaArgGly 487
Db 1579 ACCAAATGCTGCCCCCAACCAAGATTAATCAACCTTCGCTCTTCAAG----- 1632
OY 488 CysAsnProAlaProAlaProGlyCysSerProSerCysAla-----LeuThr 504
Db 1633 ACAATGCTATTTCTGATGTTCAATGTCACACCAATACAGGCAATATATATGTAAT 1692
OY 505 GlnLeuSerPheHisThrIleProThrAspSerLeuGlnTyrPheHisGlnAsnLeuGlyHis 524
Db 1693 CAAATGCTTTCACAAATAGAGATGAAGATTATTAATTAAGAAAGCTTGCGCCAA 1752
OY 525 GlySerPheThrLysIlePheArgGlySerArgArgGluValAlaAsp---GlyGlnThr 543
Db 1753 GGTACTTTTACAAAATTTTAAAGTGTAGAGAGAGAGTGGAGATTAATGCAACTG 1812
OY 544 HisAspSerGluValLeuLeuLysValMetAspSerArgHisArgAsnCysMetGluSer 563
Db 1813 CACAAACGGAGAGTCTTTTCAAGTCTAGATTAAGCACATAGGAACATTTATTCAGAGCT 1872
OY 564 PheLeuGlnAlaIleSerLeuMetSerGlnValSerTyrProHisIleValLeuLeuHis 583
Db 1873 TTCTTCAGAGACGACATGATGATGCTTTCTCACAAGCATTTGGTTTGAATATAT 1932
OY 584 GlyValCysMetAlaGly---AspSerIleMetValGlnIlePheValTyrLeuGlyAla 602
Db 1933 GGTGTGTGTCTGTGGAGAGAGAGAACTTGTGTTCAAGATTTGTAAATTTGGATCA 1992
OY 603 IleAspMetTyrLeuArgLysArgGlyHisIleValSerAlaSerTyrPheLysLeuGlnVal 622
Db 1993 CTGATACATCTCCTGAGAGAGAAACAAATTCATTAATATATTATGGAACCTTGGAGCTG 2052
OY 623 ThrLysGlnLeuAlaTyrAlaLeuAsnTyrLeuGlnAspLysGlyLeuProHisGlyAsn 642
Db 2053 GCTAAGCAGTGGCATGGCCATGATTTCTAGAGAGAAATCCCTATTATTCATGGGAT 2112
OY 643 ValSerAlaArgLysValLeuLeuAlaArgGlnGlyGly-----AspGlyAsnProPro 660
Db 2113 GTGTGTCTAAATAATCTCTTATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2172
OY 661 PheIleLysLeuSerAspProGlyValSerProThrValLeuSerLeuGlnMetLeuThr 680
Db 2173 TTTCATCAAACTTAGTACCTGGCATTTAGCATTTACAGTTTACGAAAGAGCAATCTTTCAG 2232
OY 681 AspArgTyrLeuProValAlaProGlyLysLeuGlnGlnIleGlnThrLeuCysLeuGln 700
Db 2233 GAGAGATACCATGGGTACCTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2292
OY 701 AlaAspLysTyrGlyPheGlyAlaThrThrTyrPheLysValPheGlnArgGlyProAlaHis 720
Db 2293 ACAGACAGTGGAGCTTGGAGACCACTCTGGAGATCTCAGTGGAGAGATTAAGCCC 2352
OY 721 IleThrSerLeuGluProAlaLysLysLeuLysPheTyrGlnAspGlnGlnLeuPro 740

Db 2253 CTGAGTCTCTGATCTCAAGAAACCTGAGTCTATGAGATTAAGATCACTCTCT 2412
OY 741 AlaLeuLysTyrPheThrGlyLeuAlaGlyLeuIlePheThrGlnCysMetAlaTyrAspProGly 760
Db 2413 GCACCAAGTGGAGAGAGATTAAGCAACCTTATTAATTAATTCATGATGATGATGATGATGAT 2472
OY 761 ArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIleThrSerAspTyr 780
Db 2473 TTTCAGGCTGTCTTACAGAGCTGTCATGCTGATCTTAACAGCCTGTATTACCGAGATTAT 2532
OY 781 GluLeuLeuSerAspProThrProGlyIleProSerProArg---AspGlyLeuCysVal 799
Db 2533 GAATCTTAAACAGAA---AATGACATCTACCAACATGAGAAATAGGTGCCCTTAGGCTT 2589
OY 800 AlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGlnGlnIleArgHisLeuLys 819
Db 2590 TCTGCTCT-----TTTGAAGACAGGAGCCCTACAGATTTGAAGAGACACTTGAAG 2643
OY 820 TyrIleSerLeuLeuGlyLysGlyAsnPheGlySerValGluLeuCysArgTyrAspPro 839
Db 2644 TTCTTACAGCAGCTTGGCAAAAGTAACTCGGAGTGTGGAGATGCTCCCTATGACCCG 2703
OY 840 LeuGlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGlnHisSerValProAsp 859
Db 2704 CTGAGAGCAACACTGGCGAGTGTGCTGTGAAGAACTCCGACAGCAGCTGAAGAG 2763
OY 860 GlnGlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuHisSerAspPheIle 879
Db 2764 CACCTCCGAGACTTTGAGAGAGGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 2823
OY 880 ValLysTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuArgLeuValMetGlu 899
Db 2824 GTCAAGTACCAAGGAGTGTGTACAGTGGCGGTGGCGGCAACCTTAAGATTAATTAATGAA 2883
OY 900 TyrLeuProSerGlyCysLeuArgAspLeuLeuGlnArgHisArg---GlyLeuHisThr 918
Db 2884 TATTTCATATGAGAGTTTATGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 2943
OY 919 AspArgLeuLeuLeuPheAlaTyrGlnIleCysLysGlnMetGluTyrLeuGlyAlaArg 938
Db 2944 AAAAATCTCTTCAATACATCACTCAGATATGCAAGGAGCATGAGATATCTGTGAACAA 3003
OY 939 ArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGluSerGluAlaHisVal 958
Db 3004 AGGTATATCCACAGAGAGCTTGGCAAGCAAGCAATATTTGGTGAAGAAATGACACAGGCT 3063
OY 959 LysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAspTyrTyrValAla 978
Db 3064 AAAATAGAGAGCTTGGATTAACCAAGAGCTTGGCGGACGACCAAGATTAATCAAAAGTA 3123
OY 979 ArgGluProGlyGlnSerProIlePheThrTyrAlaProGlnSerLeuSerAspAsnIle 998
Db 3124 AAGAGCGCAGGAGGAAAGCCCATATTTCTGTAGCCACACTGATCTTGCAGAGAGAGAGAG 3183
OY 999 PheSerArgGlnSerAspValTyrSerPheGlyValValLeuTyrGlnLeuPheThrTyr 1018
Db 3184 TTCTGTGGCTCAGATGATGTGGAGCTTTGGAGTGTCTATGCGAAGCTTTTACATAC 3243
OY 1019 CysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetGlyProGlnArgGlu 1038
Db 3244 ATCGAAGAGATTAAGATCCACCCCTGGAATTTATGCGAATGATTTGGCAATGATTAACAA 3303
OY 1039 GlyProProLeuCys---ArgLeuLeuGlnLeuLeuAlaGlnIleArgValArgValLeuProPro 1057
Db 3304 GGGCAAAATGATTTGTTCCATTTATGATGAGCTTACGAGAGAGAGAGAGAGAGAGAGAGAG 3363
OY 1058 ProProThrCysProThrGlnValGlnIleLeuMetGlnLeuCysTyrPheAlaProGluPro 1077
Db 3364 CCAGAGAGATCCAGATGATGATTTATGTCATGACAGAGTCTCTGACACAAATGTG 3423
OY 1078 HisAspArgProAlaPheAlaThrLeuSer 1087
Db 3424 AGCCAGGCTCCCTCTTACAGGAGACTTTG 3453

RESULT 9

US-08-665-574C-8

Sequence 8, Application US/08665574C

Patent No. 6136595

GENERAL INFORMATION:

APPLICANT: Ihle, James N.

APPLICANT: Silvenoinen, Olli

APPLICANT: Wiltuhn, Bruce A.

TITLE OF INVENTION: Jak Kinases and Regulation of Cytokine

TITLE OF INVENTION: Signal Transduction

NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS:

ADDRESS: Sterne, Kessler, Goldstein & Fox

STREET: 1100 New York Avenue, Suite 600

CITY: Washington

STATE: D.C.

COUNTRY: U.S.A.

ZIP: 20005-3934

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/665,574C

FILING DATE: 18-JUN-1996

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/282,012

FILING DATE: 29-JUL-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/097,997

FILING DATE: 29-JUL-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/118,968

FILING DATE: 09-SEP-1993

ATTORNEY/AGENT INFORMATION:

NAME: Bugalsky, Lawrence B.

REGISTRATION NUMBER: 35,086

REFERENCE/DOCKET NUMBER: 0656, 0370002/SLF/LBB

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 371-2600

TELEFAX: (202) 371-2540

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 3629 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

FEATURE:

NAME/KEY: CDS

LOCATION: 94...3480

US-08-665-574C-8

Alignment Scores:

Pred. No.: 6,47e-234

Score: 2619.50

Percent Similarity: 65.50%

Best Local Similarity: 48.38%

Query Match: 44.70%

DB: 3

US-09-397-967-16 (1-1099) x US-08-665-574C-8 (1-3629)

QY 4 ProserglutrhProleuLeuProglInarSerCysSerLeuSerSerSerGluAla 23

Db 145 CCTGTACATCAGATGATGATATTCCTGGAAGTCTCTATTCGTGAGCAATAGAGCA 204

QY 24 GAlaAlaLeuHsValLeuLeuProProArgGlyProGlyProProGlnAlaGluSerPhe 43

Db 205 GTCCCTCAAGTGTATCTGTACCAATTCCTTGCGCAAGCTGAAGGAGTATCTGAAGTTT 264

Length: 3629
 Matches: 537
 Conservative: 190
 Mismatches: 350
 Indels: 33
 Gaps: 15

QY 44 SerPheGlyAspTyrLeuAlaGluAspLeuCysValArgAlaAlaValAlaCysGlyIle 63
 Db 265 CCAAGTGGAGAGTATGTCGACAGAAATTTGTGTGCTGCTCTTAAGCTTGATGAT 324
 QY 64 LeuProValTyrHsSerLeuPheAlaLeuAlaThrGluAspPheSerCysTrpPhePro 83
 Db 325 ACGCTGTATCATATATATGTCGTTAATGACTGAACCGAAGCATCTGATGCCA 384
 QY 84 ProSerHsIlePheCysGlyGluAspValAspThrGlnValLeuValTyrArgLeuArg 103
 Db 385 CCCAATCATGCTCTCATATGACATGACGACCAACGACGATGACATCTACAGGATTAAG 444
 QY 104 PheTyrPheProAspTrpPhe-----GlyLeuGluThrCysHsArgPheGlyLeuArg 121
 Db 445 TTCTACTTCCCTCATTTGCTAGCTAGTGTGCGACAGACCAACCTACAGTACGAGTGC 504
 QY 122 LysAspLeuThrSerAlaIleLeuAspLeuHsValLeuGluHsLeuPheAlaGlnHs 141
 Db 505 CGTGGGCTGAAGCTCTCTGCTGATGACTTGTGCATGCTCTTCTTGTGCTCAGTGG 564
 QY 142 ArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLysGluGlnGlyIle 161
 Db 565 CGGCATGATTTTGTTCACAGGATGATTAAGTACCTGTGACTCATGAAGTACGAGAG 624
 QY 162 PheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGlnAlaGlnArgPro 181
 Db 625 TGCTTTGGGATGGCGGTGTAGACATGATGAAGTAAAGTAAAGGAAAGACAGCTCCA 684
 QY 182 GlyIleLeuLeuLysThrValSerTyrLysAlaCysLeuProProSerLeuArgAspVal 201
 Db 685 CTGGCTGTCTAATCTGTGACGACACAAATCTTCCAAAGTCCGCTGACGAG 744
 QY 202 IleGlnGlyGlnAspPheValThrArgArgArgIleArgArgThrValValLeuAlaLeu 221
 Db 745 ATCCAGACCTATCACAATTTTAACCGGAGGCAATCAGATTCGACAGATTCATT 804
 QY 222 LeuProCysGlyArgLeuProGlyArgProTyrAlaLeuMetAlaLysTyrIleLeuAsp 241
 Db 805 CAGCAATTCAGTCAATGTAAGCACTGCCAGAACCTTAACCTTAATCTTATTAAC 864
 QY 242 LeuGluArgLeuHsProAlaAlaThrThrGluThrPheArgVal----- 256
 Db 865 CTGGAACCTGACAGTGTGCTTCTACACAGAACGTTTGAAGTAAGAAATCTGCAGAGA 924
 QY 257 GlyLeuProGlyAlaGlnGluProGlyLeuLeuArgValAlaGlyAspAsnGlyIle 276
 Db 925 GGT---CCTTCAGGAGAGATTTTTCACACCAATTATTAATGAACGGAAGCGGATTT 981
 QY 277 ProTrpSerSer-----AsnAspGluLeuPheGlnThr 287
 Db 982 CAGTGGTCAAGAGGAAACATAGGAAGTGAACACTGACAGCAAGAGCGTACAGTTA 1041
 QY 288 PheCysAspPheProGluIleValAspValSerIleAsnGlnAlaProArgValGlyPro 307
 Db 1042 TATTGTGATTTCCCTGATATTTATGATGTCAATTAAGCAAGCAACAG---GAATGC 1098
 QY 308 AlaGlyGluHsArgLeuValThrValThrArgMetAspGlnHsIleLeuGluAlaGlu 327
 Db 1099 TCAATGAAATAGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAAT 1158
 QY 328 PheProGlyLeuProGluAlaLeuSerPheValAlaLeuValAspGlyTyrPheArgLeu 347
 Db 1159 CTTAGCTCATTTAAAGAGCCCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1218
 QY 348 IleCysAspSerArgHsIleTyrPheCysLysGluValAlaProProArgLeuGlnGlu 367
 Db 1219 ACTCGGATGCGCACCACTTACCTTGCAGAGAGGTGGCTCCCGACGCTGTGAGAAC 1278
 QY 368 GluAlaAspValCysHsArgProIleThrLeuAspPheAlaIleHsLysLeuAla 387
 Db 1279 ATACACAGCAACTGCGACGCGCAATGATGATGATTTTGGCAATTAACAAAGAG 1338
 QY 388 AlaGlySerLeuProGlyThrTyrIleLeuArgArgSerProGlnAspTyrAspSerPhe 407

1339 GCGGGTACACAGACTGACACTATATGCTACGATGCGCCCTTAAGCCTTCAACAAATAC 1398
408 LeuLeuThrAlaCysValGlnThrProLeuGlyProAspTyrLysGlyCysLeuIleArg 427
1399 TTCTGACCTTTGGCTGGTGGAGCGAAGAAATGTCATGAATATAAACACTGTTGATTACG 1458
428 GlnAspProSerGlyAlaPheSerLeuValGlyLeuSerGlnProHisAsnSerLeuArg 447
1459 AAGATAGAGATGAGAGATCAACCTCAGCGGACCTAAGAGGAACCTTCAGTAACCTTAAG 1518
448 GlnLeuLeuAlaIaCysTrpAsnSerGlyLeuArgValAspGlyAlaAlaLeuTyrLeu 1578
1519 GACCTTTTGATTTGCTACGACGATGGAACCTGCGCTCAGACGATACATCTTCCACTTT 1578
468 ThrSerCysAlaProArgProLysGlyLysSerAsnLeuIleValValArgArgGly 487
1579 ACCAAATGCTGCCGCCCAAGGCCAAAGATAAATCAAACTTCTCGCTTCAGA----- 1632
488 CysAsnProAlaProAlaProGlyCysSerProSerCysAla-----LeuThr 504
1633 ACAATGCTATTCTGATGTTGATCTCAGATCTCAGCAACATTACAGAGCGCATATATGTAAT 1692
505 GlnLeuSerPheHisThrIleProThrAspSerLeuGlnTyrPheHisGlnAsnLeuGlyHis 524
1693 CAATGGTGTTCACAAATCAGACGATGAAGATTAATATTAATGAAGACTGTGGCCAA 1752
525 GlySerPheThrLysIlePheArgGlySerArgArgGlyValAlaLysPheGlyGlnThr 543
1753 GGTCTTTTACAAAAATTTTAAAGGTGTAAAGAGAGAGATGAGATTAATGATGCAACTG 1812
544 HisAspSerGluValLeuLeuLysValMetAspSerArgHisArgAsnCysMetGluSer 563
1813 CACAAACGAGAGTCTTTTCAAAAGCTCTAGATTAAGACACATAGACACTATTCAGAGCT 1872
564 PheLeuGlnAlaAlaSerLeuMetSerGlnValSerTyrProHisLeuValLeuLeuHis 583
1873 TTCTTCAGACGACGACACATGATGATGATGCTTTCACAAAGCATTTGGTTTAAATAT 1932
584 GlyValCysMetIaGlyLysAspSerIleMetValGlnGlnPheValTyrLeuGlyVal 602
1933 GGTCTGCTGCTGCTGGAGAGGAGACATTTGCTGTTCAAGATTTGTTAAATTTGGATCA 1992
603 IleAspMetTyrLeuArgLysArgGlyHisLeuValSerAlaSerTyrLysLeuGlnVal 622
1993 CTGATGATCATCTGAGAGAACAAATTCATAAATATATATATGTAAGACTTGGAGAT 2052
623 ThrLysGlnLeuAlaTyrAlaLeuAsnTyrLeuGlnAspLysGlyLeuProHisGlyAsn 642
2053 GCTAAGAGATGGATGGCATGGCCATGTCATTTCTAAGAAATAATCCCTTATTCATGGGAT 2112
643 ValSerAlaArgLysValLeuLeuAlaArgLysGlyLysPheValAsnProPro 660
2113 GTGTGCTGCTAAATATTCCTCTTATCAGAGACAGACAGACAGAGAGAGAGAGAGAGAGAG 2172
661 PheIleLysLeuSerAspProGlyValSerProThrValLeuSerLeuGlnMetLeuThr 680
2173 TTTCATCAAACTTAGTGAATCCGCGATTCATGATGATTCATGCGAAGAGAGAGAGAGAGAG 2232
681 AspArgLysProTyrValAlaProGlnCysLeuGlnGlnIaGlnThrLeuCysLeuGln 700
2233 GAGGAAATACCATGGGTAACCTCGAATGCAATGAGAAATCCTTAAATCTCAATCTGAGCA 2292
701 AlaAspLysTrpGlyPheGlyAlaIaThrThrTrpGlnValPheGlnArgGlyProAlaHis 720
2293 ACAGACAGATGGAGCTTGGGAGACACTCTGTGGAGATCTGACAGGAGAGAGATAGAGCC 2352
721 IleThrSerLeuGlnProAlaLysLysLeuLysPheTyrGlnAspGlnGlyLeuPro 740
2353 CTGAGTGTCTGGAATTCACAAAGAAAGCTGAGTTCTATGAAGATTAACATCATCTTCTCT 2412
741 AlaLeuLysTrpThrGlnLeuAlaGlyLeuIleThrGlnCysMetAlaTyrAspProGly 760

2413 GCACCCAGTGGACAGAGTAGCAAACTTATAAATATTCATGAGCTATAGAGCCAGAT 2472
761 ArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIleThrSerAspTyr 780
2473 TTCAGGCTCTCTTTCAGAGCTGATCCGATCCGATCTTAACAGCCTGTTTACTCCAGATTAT 2532
781 GlnLeuLeuSerAspProThrProGlyIleProSerProArgLysPheVal 799
2533 GAACCTACTACAGAA---AATGACATGCTACCAAACTAGAAATAGGTGCTTAGGGTTT 2589
800 AlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGlnGlnArgHisLeuLys 819
2590 TCTGTTGCT-----TTTGAAGACAGGAGACCTTACACAGTGTGAAGAGACACTTGAAG 2643
820 TyrIleSerLeuLeuGlyLysGlyAsnPheGlySerValGlnLeuCysArgTyrAspPro 839
2644 TTCTTCAGCAGCTTGGCAAGGTAACCTTGGAGTGTGAGATGTGCTTATGACCCG 2703
840 LeuGlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGlnHisSerValProAsp 859
2704 CTGAGAGCAACACTGGCGAGGTGTGCTGTGAAGAACTCCAGACAGACTGAAAGAG 2763
860 GlnGlnArgAspPheGlnArgGlnIleGlnIleLeuLysAlaLeuHisSerAspPheIle 879
2764 CACCTCGAGACTTGGAGAGGATGAGATGATCTGTAATCTTGCAGCATGCAACATC 2823
880 ValLysTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuArgLeuValMetGlu 899
2824 GTCAAGTACAAAGGAGAGTGTCTACAGTGGCGGTGGCGGCAACCTAAGATTAATATGAA 2883
900 TyrLeuProSerGlyCysLeuArgAspLeuLeuGlnArgHisArgLysGlyLeuHisThr 918
2884 TATTTACCATATGGAAGTGTACGAGACTATCTCCAAATGAAGAAAGAGATGATCAC 2943
919 AspArgLeuLeuLeuPheAlaTrpGlnIleCysLysGlyMetGlnTyrLeuGlyAlaArg 938
2944 AAAAATCTCTTCATATACATCTCAGATTCGCAAGGCGATGGAATTCCTGGTACAAAA 3003
939 ArgCysValHisArgAspLeuAlaIaArgAsnIleLeuValGlnSerGluAlaHisVal 958
3004 AGTATATTCACAGAGGAGCTGGCAACAAAGACATATTTGGTGGAAATGAGAACAGGTT 3063
959 LysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAspTyrTyrValVal 978
3064 AAAATAGAGAGCTTGGATTAACCAAGTCTTGGCGAGCAAGAAAGATCTACAAAGTA 3123
979 ArgGlnProGlnSerProIlePheTrpTyrAlaProGlnSerLeuSerAspAsnIle 998
3124 AAGAGCGCAGGAGAAACCCCATATTTCTGTAGCCAGCTGATCTTGACGAGAGAGAG 3183
999 PheSerArgGlnSerAspValTrpSerPheGlyValValLeuTyrGlnLeuPheThrTyr 1018
3184 TTCTCTGCGCTCAGATGATGTTGGAGCTTGGAGTGTCTATACGAACCTTTCACATAC 3243
1019 CysAspLysSerCysSerProSerAlaGlnPheLeuArgMetMetGlyProGlnLysGln 1038
3244 ATCGAAGAGATAAAGTCCACCCGTGGAATTTATGCGAATGATGGCAATGATAAACA 3303
1039 GlyProProLeuCysLysArgLeuGlnLeuLeuAlaGlnLysArgArgLeuProPro 1057
3304 GGGCAATGATTTGTTCCATTTGATAGAGCTACTGAAGACAGCAAGAGATTCACAAAG 3363
1058 ProProThrCysProThrGlnValGlnGlnLeuMetGlnLeuCysTrpAlaProGlnPro 1077
3364 CCAGAAAGATCCCAAGATTAATGATCATGTACAGACAGTCTGACAGCAACAAATGTG 3423
1078 HisAspArgProAlaPheAlaThrLeuSer 1087
3424 AGCCAGGCTCCCTTCAGGAGACTTTCG 3453
RESULT 10
US-08-946-994-8
; Sequence 8, Application US/08946994


```

: Patent No. 6210654
:
: GENERAL INFORMATION:
:
: APPLICANT: Title, James N.
:
: APPLICANT: Silvennoinen, Olli
:
: APPLICANT: Wiltuhin, Bruce A.
:
: APPLICANT: Queller, Frederick W.
:
: TITLE OF INVENTION: Jak Kinases and Regulation of Cytokine Signal
:
: TITLE OF INVENTION: Transduction
:
: NUMBER OF SEQUENCES: 17
:
: CORRESPONDENCE ADDRESS:
:
: ADDRESSEE: Sterne, Kessler, Goldstein & Fox
:
: STREET: 1100 New York Avenue, Suite 600
:
: CITY: Washington
:
: STATE: D.C.
:
: COUNTRY: U.S.A.
:
: ZIP: 20005-3934
:
: COMPUTER READABLE FORM:
:
: MEDIUM TYPE: Floppy disk
:
: COMPUTER: IBM PC compatible
:
: OPERATING SYSTEM: PC-DOS/MS-DOS
:
: SOFTWARE: Patentin Release #1.0, Version #1.25
:
: CURRENT APPLICATION DATA:
:
: APPLICATION NUMBER: US/08/946,994
:
: FILING DATE:
:
: CLASSIFICATION:
:
: PRIOR APPLICATION DATA:
:
: APPLICATION NUMBER: US/08/665,574
:
: FILING DATE: 18-JUN-1996
:
: APPLICATION NUMBER: 08/282,012
:
: FILING DATE: 29-JUL-1994
:
: PRIOR APPLICATION DATA:
:
: APPLICATION NUMBER: 08/097,997
:
: FILING DATE: 29-JUL-1993
:
: PRIOR APPLICATION DATA:
:
: APPLICATION NUMBER: 08/118,968
:
: FILING DATE: 09-SEP-1993
:
: ATTORNEY/AGENT INFORMATION:
:
: NAME: Fox, Samuel L.
:
: REGISTRATION NUMBER: 30,353
:
: REFERENCE/DOCKET NUMBER: 0656.0370002/SLF/GKT
:
: TELECOMMUNICATION INFORMATION:
:
: TELEPHONE: (202) 371-2600
:
: TELEFAX: (202) 371-2540
:
: INFORMATION FOR SEQ ID NO: 8:
:
: SEQUENCE CHARACTERISTICS:
:
: LENGTH: 3629 base pairs
:
: TYPE: nucleic acid
:
: STRANDEDNESS: single
:
: TOPOLOGY: linear
:
: FEATURE:
:
: NAME/KEY: CDS
:
: LOCATION: 94..3480
:
: US-08-946-994-8
:
: Alignment Scores:
:
: Pred. No.: 6.47e-234 Length: 3629
:
: Score: 2619.50 Matches: 537
:
: Percent Similarity: 65.50% Conservative: 190
:
: Best Local Similarity: 48.38% Mismatches: 150
:
: Query Match: 44.70% Indels: 33
:
: DB: 4 Gaps: 15
:
: US-09-397-967-16 (1-1099) x US-08-946-994-8 (1-3629)
:
: Oy 4 ProSerGluGluThrProLeuIleProGlnArgSerCysSerLeuSerSerGlnAla 23
: ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
: Db 145 CCTGACATCTGAGATGCGATATTCCTCGGAAGTCTAATTCGTGGAACAGATGACCA 204
:
: Oy 24 GlyAlaLeuHisValLeuLeuProProArgGlyProGlyProGlyProGlnArgLeuSerPhe 43
: ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
: Db 205 GTCTTCAAGTGTATCTGTACCATTCCTTCTGGCAGACGTGAAGAGAGTATCTGAAGTTT 264
:
: Oy 44 SerPheGlyAspArgTyrLeuAlaGlnAlaSerLeuCysValArgAlaAlaValAlaCysGlyIle 63

```

Db 1339 GCGGGTAACAGACTGACATATATGTGCTACGATCGAGCCCTAAGGACTTCAACAAATAC 1398
OY LeuLeuThrAlaCysValGlnThrProLeuGlyProAspTyrIleuGlyLeuValLeuArg 427
Db 1399 TTTCTGACCTTCTGCTGTGTGCGGAGAAATGTCATTGATTAACACGTTTGTGATTACG 1458
OY 428 GlnAspProSerGlyAlaPheSerLeuValGlyLeuSerGlnProHisArgSerLeuArg 447
Db 1459 AAGATGACAAATGAGAAATACAACTCAGGAGACTAAGAGAACTTCAGTAACTTAAG 1518
OY 448 GlnLeuLeuAlaIleCysTyrPasnSerGlyLeuArgValAspGlyAlaAlaLeuTyrLeu 467
Db 1519 GACCTTTGAAATGCTCAGACATGAAACTGCTGCTCAACAGATATCATCTTCCAGTTT 1578
OY 468 ThrSerCysAlaProArgProGlyGlnLeuSerAsnLeuLeuValAlaArgArgGly 487
Db 1579 ACCAAATGCTGCCCCCAAGGCCAAAGATTAATCAACCTTCTGCTTCACA----- 1632
OY 488 CysAsnProAlaProAlaProGlyCysSerProSerCysCysAla-----LeuThr 504
Db 1633 ACAATGCTATTTCTGATGTTCAAGATCTCAACATTAACAGGCAATTAATGTGAAT 1692
OY 505 GlnLeuSerPheHisThrIleProThrAspSerLeuGlnTyrPheHisGlnLeuGlnHis 524
Db 1693 CAAATGCTTTCACAAATCAGCAATGAAATTAATTAATTAATGAAGCTTGGCCAA 1752
OY 525 GlySerPheThrIlePheArgGlySerArgGlyValAlaAsp---GlyGluThr 543
Db 1753 GGTACTTTACAAAAATTTTAAAGGTGTAAAGAGAGAGAGTGAAGTATGTGCAACTG 1812
OY 544 HisAspSerGlnValLeuLeuValMetAspSerArgHisArgAsnCysMetGlnSer 563
Db 1813 CACAAAGCGAAGTCTTTGAAAGTCTTACATTAAGCACATAGCACTATTCAGACTCT 1872
OY 564 PheLeuGlnAlaAlaSerLeuMetSerGlnValSerTyrProHisLeuValLeuLeuHis 583
Db 1873 TTTCTGCAAGCAGCAGCATGATGATGCTTCTCACAAGCATTTGCTTTAAATAT 1932
OY 584 GlnValCysMetAlaGly---AspSerIleMetValGlnGlnPheValTyrLeuGlyAla 602
Db 1933 GGTGCTGTGCTGTGAGAGAGAGACATCTGCTTCAAGAAATTTGTAAATTTGGATCA 1992
OY 603 IleAspMetTyrLeuArgIleValArgGlyHisLeuValSerAlaSerTyrPheLeuGlnVal 622
Db 1993 CTGGATATACCTCAGACAGAAACAAAAATTCATTAATATTAATGAAATCTGGAGATG 2052
OY 623 ThrIleGlnLeuAlaTyrAlaLeuAsnTyrLeuGlnAspIleGlyLeuProHisGlyAsn 642
Db 2053 GCTAAGCAGATTGGCATGGCCATGCTATTCAGAAAGAAATCCCTTATTCATGGGAT 2112
OY 643 ValSerAlaArgIleValLeuLeuAlaArgGlnGlyGly-----AspGlyAsnProPro 660
Db 2113 GGTGTGCTAAATAATCTCTGCTTATCAGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAG 2172
OY 661 PheIleIleuSerAspProGlyValSerProThrValLeuSerLeuGlnMetLeuThr 680
Db 2173 TTTCACTAACTAGTATCTGTCATTCAGATTACAGTTTCAGGAGACATTTCTTAG 2232
OY 681 AspArgIleProThrValAlaProGlnCysLeuGlnGlnAlaGlnIleThrCysLeuGln 700
Db 2233 GAGAGAAATCCATGGTACTCTGATGCAATGCAATGCAATGCAATGCAATGCAATGCAATG 2292
OY 701 AlaAspIleTyrPheGlyAlaThrThrTyrPheGlyValAlaPheGlnArgIleProAlaHis 720
Db 2293 ACAGACAG 2352
OY 721 IleThrSerLeuGlnProAlaIleCysLeuIleuSerPheTyrGlnAspGlnGlnLeuPro 740
Db 2353 CTGAGTGTCTCTGATTCCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2412
OY 741 AlaLeuIleTyrThrGlnLeuAlaGlyLeuIleThrGlnCysMetAlaTyrAspProGly 760
Db 2413 GCACCCAGAGTGCAGAGATTAGCAAACTTAAATAATTAATTAATTAATTAATTAATTAAT 2472

OY 761 ArgArgProSerPheAlaValAlaIleLeuArgAspLeuAsnGlyLeuIleThrSerAspTyr 780
Db 2473 TTTACGCTGCTTTCAGAGCTGTGATCGTATCTTAACAGCCTGTTTACTCCAGATTAAT 2532
OY 781 GlnLeuLeuSerAspProThrProGlyIleProSerProArg---AspGlnLeuCysVal 799
Db 2533 GAACTACTTAACAGAA---AATGACATGCTTACCAAAATGAGAAATAGAGAGAGAGAGAG 2589
OY 800 AlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGlnGlnGlnGlnGlnGln 819
Db 2590 TCTGCTGCT-----TTTGAAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2643
OY 820 TyrIleSerLeuLeuGlnGlyLeuGlnAsnPheGlySerValGlnLeuCysArgTyrAspPro 839
Db 2644 TTTTCTACAGCAGCTTGGCAAGAGTAACTTCCGAGAGTGGAGAGATGCGCTATGACCG 2703
OY 840 LeuGlyAspAsnThrGlyProLeuValAlaValIleGlnLeuGlnHisSerValProAsp 859
Db 2704 CTGACAGACACACACTGGCAGAGTGTGCTGTGAAGAAATCCACAGCAGCAGAGAGAGAG 2763
OY 860 GlnGlnArgAspPheGlnArgGlnIleGlnIleLeuValAlaLeuHisSerAspPheIle 879
Db 2764 CACTTCCGAGCTTTCAGAGAGAGATCCAGATCTCGAAATCTTCCAGCAGCAGAGAGAG 2823
OY 880 ValIleTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuArgLeuValMetGlu 899
Db 2824 GTCAAGTACAAAGAGAGAGTGTCTACAGTGTGCGGCTGCGGCAACCTTAATTAATTTGGA 2883
OY 900 TyrLeuProSerGlyCysIleuArgAspLeuGlnArgHisArg---GlyLeuIleThr 918
Db 2884 TATTTACATGAGAGAGTTCAGAGACTTCTCCAAAACATTAAGAGAGAGAGATATCAG 2943
OY 919 AspArgLeuLeuLeuPheAlaThrGlnIleCysIleGlyMetGlnTyrLeuGlyAlaArg 938
Db 2944 AAAAATCTTCAATACATACATCATCTCAGATATGCAAGAGCAGATATCTTGTGTCAAA 3003
OY 939 ArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGlnSerGlnAlaHisVal 958
Db 3004 AGGTATATCCAGAGAGAGCTGCGACACAGAGACATTTGTGGAATATGAGAACAGGTT 3063
OY 959 LysIleAlaAspPheGlyLeuAlaIleValIleLeuProLeuGlyLysAspTyrTyrValVal 978
Db 3064 AAAATAGAGACCTTGGATTAACCAAAAGCTTCCAGAGACAAAGATTAACAAAGTA 3123
OY 979 ArgGlnProGlyGlnSerProIlePheTyrPyrAlaProGlnSerLeuSerAsnIle 998
Db 3124 AAGGAGCCAGAGGAGAAAGCCCATATCTGTGTACGCACTGAATCTTGAAGAGAGAGAG 3183
OY 999 PheSerArgIleSerAspValIleTyrPheGlyValValLeuTyrGlnLeuPheThrTyr 1018
Db 3184 TTTTCTGTGCTCAGATGTGTGTGAGCTTTGAGTGTCTTATACGACTTTTACATATAC 3243
OY 1019 CysAspIleSerCysSerProSerAlaGlnPheLeuArgMetMetGlyProGlnAlaArgGlu 1038
Db 3244 ATCAGAGAAAGATTAAGCTCCACCCGTGGAATTTATGTGAATGTGATGCAATGAACA 3303
OY 1039 GlyProProLeuCys---ArgLeuLeuGlnLeuLeuAlaGlnIleArgArgLeuProPro 1057
Db 3304 GGGCAAAATGATGTCTTCATTTATATAGACTACTGAAAGAGCAAGAGAGAGAGAGAGAG 3363
OY 1058 ProProThrCysProThrGlnValGlnGlnLeuMetGlnLeuCysTyrPheAlaProGlnPro 1077
Db 3364 CCAGAGAGATCCCAAGATGATTAATGTATGTATGTATGTATGTATGTATGTATGTATGT 3423
OY 1078 HisAspArgProAlaPheAlaThrLeuSer 1087
Db 3424 AGCCAGCGTCCCTCTTCAGGAGCCTTTCG 3453

RESULT 11
US-08-446-038B-2
; Sequence 2, Application US/08446038B
; Patent No. 5658791

```

: GENERAL INFORMATION:
: APPLICANT: Wilks, Andrew F.; Ziemiecki, Andrew;
: APPLICANT: Harpur, Ailsa
: TITLE OF INVENTION: No. 5658791el Protein Tyrosine Kinase
: NUMBER OF SEQUENCES: 23
: CORRESPONDENCE ADDRESSES:
: ADDRESS: Felte & Lynch
: STREET: 805 Third Avenue
: CITY: New York City
: STATE: New York
: COUNTRY: USA
: ZIP: 10022
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
: COMPUTER: IBM PS/2
: OPERATING SYSTEM: PC-DOS
: SOFTWARE: Wordperfect
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/446.038B
: FILING DATE: 19-MAY-1995
: CLASSIFICATION: 424
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/064,067
: FILING DATE: 30-Jun-1993
: APPLICATION NUMBER: PCT/US91/08889
: FILING DATE: 26-No. 5658791-1991
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: Australian PK3594/90
: FILING DATE: 28-No. 5658791-1990
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: Australian 88229/91
: FILING DATE: 27-No. 5658791-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: Hanson, No. 5658791man D.
: REGISTRATION NUMBER: 30,946
: REFERENCE/DOCKET NUMBER: LUD 5244
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 212-688-9200
: TELEFAX: 212-638-3884
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3495 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: nucleic acid
: US-08-446-038B-2

```

Alignment Scores:

Pred. No.:	9,43e-210	Length:	3495
Score:	2360.00	Matches:	489
Percent Similarity:	65.89%	Conservative:	164
Best Local Similarity:	49.34%	Mismatches:	300
Query Match:	40.27%	Indels:	38
DB:	1	Gaps:	15

```

US-09-397-967-16 (1-1099) x US-08-446-038B-2 (1-3495)
QY 128 lleleAspIeuHIsValleuHIsleuPheAlaGlnHIsarGerAspIeuValSer 147
   ::::::::::::::::::::
Db 1 CTGCTTGATGACTTCTTCATGCTTACCTTCCCTCAGTGGCGGATGATTTTGTTCAC 60
QY 148 GluArgIeuProValGlyIeuSerMetIySgInGlnGlyIeuPheIeuSerIeuAlaVal 167
   ::::::::::::::::::::
Db 61 GGATGGATGAAGTACCTGCTGCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
QY 168 leuAspIeuAlaGlnMetAlaIarGlnGlnAlaGlnArgProGlyIeuIeuLeuIyStr 187
   ::::::::::::::::::::
Db 121 TTAGACATGATGAGATTAAGTAAAGGAGAAAGACCAAGTCCAGCTGCTGTATTAAGTCT 180
QY 188 ValSerTylIyAlaCysIeuProPheSerIeuArgAspValIleGlnGlnIasnPhe 207
   ::::::::::::::::::::
Db 181 GTCAGCTACAGACATTTCTTACCAGGAGGCTTTCGAGGAGAGATCCAGACATTCACATT 240

```

```

QY 208 ValThrArgArgIleArgArgThrValValleuAlaIleuProCysIleArgIeu 227
   ::::::::::::::::::::
Db 241 TTAACCCGGAGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
QY 228 ProGlyArgProTyrAlaIeuMetAlaIyTyrIleIeuAspIeuGlnArgIeuHIsPro 247
   ::::::::::::::::::::
Db 301 AAGCCACTGCCAGGAACTTAACCTTAAGTATCTTTAAGCTGGAACCCCTGAGTCT 360
QY 248 AlaAlaThrThrGluThrPheArgVal-----GlyIeuProGlyAlaGln 262
   ::::::::::::::::::::
Db 361 GCCCTTCTACACAGACAGTTCGATTAAGATTAAGATTCGAAAGAGT---CCTTCAGGTGAG 417
QY 263 GluIeuProGlyIeuIeuArgValAlaGlyAspAsnGlyIleProIyPheSer----- 280
   ::::::::::::::::::::
Db 418 GAGATTTTTCGAAACCATTAATTAAGTGAAGAACCGGTGAATTCAGGTGCAAGAGGANA 477
QY 281 -----AsnAspGluIeuPheGlnThrPheCysAspIeuProGlu 293
   ::::::::::::::::::::
Db 478 CATAGGAAAGTGAAGACTGACAGAACAGAGAGTATTAATGATGATTTCCCTGAT 537
QY 294 lleValAspValSerIleAsnGlnAlaProArgValGlyProAlaGlyIeuHIsArgIeu 313
   ::::::::::::::::::::
Db 538 ATTATGATGATGATGATTAAGCAAGCAATCAG---GAATGCTCACTGAAGTAAAGT 594
QY 314 ValThrValThrArgMetAspGlnHIsIleIeuGlnAlaGluPheProGlyIeuProGlu 333
   ::::::::::::::::::::
Db 595 GTGACCGTCCACAGAGAGCGGAGAGGCTTGGAATTAAGAACTGATGATGATTAAGAA 654
QY 334 AlaIeuSerPheValAlaIeuValAspGlyTyrPheArgIeuIleCysAspSerArgHIs 353
   ::::::::::::::::::::
Db 655 GCCCTGTCATGCTGATTAATTAAGTGAAGGATTAAGTGAAGTGAAGTGAAGTGAAG 714
QY 354 TyrPheCysIleGlnValAlaProProArgIeuGlnGluIeuAlaAspValCysHIs 373
   ::::::::::::::::::::
Db 715 TACCTTGCAGAAAGAGTGGCTGCCAGAGTGTGTGAGAACATACAGCAACTGCCAC 774
QY 374 GlyProIleThrIeuAspPheAlaIleHIsIeuIySgInGlnGlnAlaIleSerIeuProGly 393
   ::::::::::::::::::::
Db 775 GCCCAATTTCAATGATTTTGCATCAGCAACTTAAGAGAGCAAGCAAGCAAGCAAGTGA 834
QY 394 ThrTyrIleIeuArgArgSerProGlnAspTyrAspSerPheIeuIeuThrIleCysVal 413
   ::::::::::::::::::::
Db 835 CTGTATGTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 894
QY 414 GlnThrProIeuGlyProAspTyrIySgInGlnGlnIleArgIeuAspProSerGlyAla 433
   ::::::::::::::::::::
Db 895 GAGCGGAAAGTGTATTGATTAATTAACACTGTTGATTAACAAGATGAGATGGAGAG 954
QY 434 PheSerIeuValGlyIeuSerGlnProHIsArgSerIeuArgIleIeuIeuAlaIaCys 453
   ::::::::::::::::::::
Db 955 TACAACTGAGGAGGAGTAAAGAGAACTTCAGAGTCTTAAGAGAACTTTGAATTTGCTAC 1014
QY 454 TrpAsnSerGlyIeuArgValAspGlyAlaIleIeuTyrIeuIeuSerCysAlaIlePro 473
   ::::::::::::::::::::
Db 1015 CAGATGGAAGTGTGGCTGACAGACTGATCCTTCCAGTTCACCAAAAGTCTCTCTCCA 1074
QY 474 ArgProIySgInGlnIySgInSerIeuIleValIleArgArgGlyCysAsnProAlaProAla 493
   ::::::::::::::::::::
Db 1075 AAGCGGAAAGATTAATCAAACTTCTGCTTGA---ACAATAGTGTGTTGCTGAT 1128
QY 494 ProGlyCysSerProSerCysAla-----LeuThrGlnIeuSerPheHIsThr 510
   ::::::::::::::::::::
Db 1129 GTTCACCTCTCACCAACTTACAGAGCATTAATTAATGATGATTAATGATGATGATGAT 1188
QY 511 IleProThrAspSerIeuGlnTrpHIsGlnIeuGlnIySgInSerPheThrIySle 530
   ::::::::::::::::::::
Db 1189 ATCAGCAATGAGATTTGATTAATTAAGAGCTTGGCCAGAGCACTTTTACAAATA 1248
QY 531 PheArgGlySerArgArgIeuValIleAsp---GlyIeuThrHIsAspSerGlyIeuIleu 549
   ::::::::::::::::::::
Db 1249 TTTAAAGGTGTAAGAGAGAGTGAAGTATTAATGATGATGATGATGATGATGATGATGAT 1308

```

QY	550	LeIysValIeMLeTAserATGHSIAsrASrASrCysMeTGlSerPheLeuGluAlaIleAser	569
	
Db	1309	TTGAAAGCTCTAGATTAAGACACATAGAAACTATTCACAGACTCTTTCTTGAAACACACAGC	1368
QY	570	LeuMetSerGlnValSerTyrProHisLeuValLeuLeuHisGlyValCysMetAlaGly	589
	
Db	1369	ATGATGAGTACAGCTTCTCCACAGACATTTGGTTTTGAATTAATGAGAGTATGTCTGTCTGGA	1428
QY	590	---AspSerIleMetValGlnGluPheValTyrLeuGlyAlaIleAspMetTyrLeuArg	608
	
Db	1429	GAGGAGAACATTTTGCTTCAGAGATTGTAAATTTTGATCCTGATGCATACACTCTGAG	1488
QY	609	LysArgGlyHisLeuValSerAlaSerTyrLeuGlnValThrGlySerLeuValLys	628
	
Db	1489	AAGAACAAAAATTTCTATTAATATATTATGTGAACACTTCGAGTGGCCGACAGACTTGGCATGG	1548
QY	629	AlaLeuAsnTyrLeuGlnAspLysGlyLeuProHisGlyAsnValSerAlaArgVal	648
	
Db	1549	GCCATGCACTTCTCTCGAAGAAAAATCCCTTATTCATGGAAATGTGTGCTTAATAATATC	1608
QY	649	LeuLeuAlaArgGluGlyGly-----AspGlyAsnProPheIleLysLeuSerSp	666
	
Db	1609	CTGCTTATCAGAGAAAGACACAGAGAACGGGAGAACCCACTTTCATCAAACTATAGTAT	1668
QY	667	ProGlyValSerProThrValLeu-----SerLeuGlnMetLeu	679
	
Db	1669	CTGTGCATTACCATTAAGTCTTACCGAAGACATTTCTCTGTGTGTTCCAACTCTT	1728
QY	680	ThrAspArgIleProTyrValAlaProGlyCysLeuGlnGluAlaGlnThrLeuCysLeu	699
	
Db	1729	CAGGAGAAATACCATCGGTATCCACCCGTGATGTCATTGAATCTCAAAATCTTACCTGTG	1788
QY	700	GluAlaAspLysTyrGlyPheGlyAlaThrThrTrpGluValPheGlnArgGlyProAla	719
	
Db	1789	GCAACAGACAAAGTGACCTTCGGACCACTCTGTGGAAATCTTGACATGGAGACATAG	1848
QY	720	HisIleThrSerLeuGluProAlaLysLeuLysPheTyrGluAspGlnGlyLeu	739
	
Db	1849	CCCTGAGTGTCTGATCTCTCAAGAAAGCTGCATCTTATGAAAGATACATCAGCTT	1908
QY	740	ProAlaLeuLysTyrThrGluLeuAlaGlyLeuIlePheGlnCysMetAlaTyrAspPro	759
	
Db	1909	CCCTGACCCAAAGTGCAGACAGATTGGCAAACTTATTAATATTCATGACATGACACCA	1968
QY	760	GlyArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIlePheSerSp	779
	
Db	1969	GATTTCAGGCTGCTTTCACAGCTGTATCCGGATCTTAAACGCTCTTATCTCAAT	2028
QY	780	TyrGlnLeuLeuSerAspProThrProGlyIleProSerProArg---AspGluLeuCys	798
	
Db	2029	TATGACTACTACAAACGAA---AATGACATGCTACCAACATAGAAATAGTGCCTTAGG	2085
QY	799	ValAlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGlnGluArgHisLeu	818
	
Db	2086	TTTTCTGTGCTCT-----TTTGAAGACAGAGGACCTTACACATTTGAAGAGACACTTG	2139
QY	819	LysTyrIleSerLeuLeuGlyLysGlyAsnPheGlySerValGluLeuCysArgTyrAsp	838
	
Db	2140	AAGTTTCTACAGCAGCTTGGCAAAAGTAATCTTCGGAGTGTGAATGTGCTCCGCTATGAC	2199
QY	839	ProLeuGlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGlnHisSerValPro	858
	
Db	2200	CCGCTGAGGACAACTGGCGAGGTGTGCTGTAAAGAACTCCACACAGCATGAA	2259
QY	859	AspGlnGlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuHisSerAspPhe	878
	
Db	2260	GAGACCTCCGAGACTTTTGAGAGGGAGATCGAGATCTCTAAATCTCTTGACAGATGACAC	2319
QY	879	IleValLysTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuArgLeuValMet	898
	
Db	2320	ATGCTCAAGTACAAAGGAGGTGCTCTACAGTGGGGGTGGCGCAACCTTAAGATTAAATATG	2379
QY	899	GluTyrIleuProSerGlyCysLeuArgAspLeuLeuGlnArgHisArg---GlyLeuHis	917


```

Oy 1038 GtAluIyProProLeucCys---ArgLeuLeuGluLeuLeuAlaGluGlyArgArgLeuPro 1056
      ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 2800 CAAGGCAAAATGATCTGTTCCTCATTTGATATAGAGCTACTAGAGCAACGAGATTGCCA 2858
Oy 1057 ProProProthCysProThrGluValGlnGluLeuMetGlnLeuCysTrpAlaProGlu 1076
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2860 AGCCGACAGAGATGCCACAGATGAGATTTCATGATCATGACAGAGAGTGTGAACAACAAT 2919
Oy 1077 ProHisAspArgProAlaPheAlaThrLeuSer 1087
      ||||| ::::: |||||
Db 2920 GTGAGCCAGCGTCCCTCCTTCAGGAGCACTTTC 2952

RESULT 13
US-08-805-445-2
Sequence 2, Application US/08805445
Patent No. 5821069
GENERAL INFORMATION:
APPLICANT: Wilks, Andrew F.; Ziemiecki, Andrew;
APPLICANT: Harpur, Ailsa
TITLE OF INVENTION: No. 5821069el protein Tyrosine Kinase
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Felfe & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/805,445
FILING DATE: 25-FEB-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/446,038
FILING DATE: 19-MAY-1995
APPLICATION NUMBER: 08/064,067
FILING DATE: 30-Jun-1993
APPLICATION NUMBER: PCI/US91/08889
FILING DATE: 26-No. 5821069-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: Australian PK3594/90
FILING DATE: 28-No. 5821069-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: Australian 88229/91
FILING DATE: 27-No. 5821069-1991
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 5821069man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 5244
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-688-9200
TELEFAX: 212-838-3884
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 3495 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: nucleic acid
US-08-805-445-2

Alignment Scores:
Pred. No.: 9.43e-210 Length: 3495
Score: 2360.00 Matches: 489
Percent Similarity: 65.89% Conservative: 164
Best Match Similarity: 49.34% Mismatches: 300
Query Match: 40.27% Indels: 38

```



```

Db      2086 TTTTCGTGCTCT-----TTTGAACAGCAGGACCTTACACAGTTTGAAGAGACACTTG 2139
Oy      819 LYSTYLILeserleuenglyLysGlyAsnphleglySerValGluLeucysarGlyrAsp 838
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      2140 AAGTTTCTACAGACGCTTGCAAAAGCTAACTTCGGAGAGTGTGAGATGTGCCGTATGAC 2199
Oy      839 PROleuglyspasnThrGlyProleuValAlaValLysGluLeuGlnHisSerValPro 858
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      2200 CCGCTGAGAGCAACACTGCGAGGTGGTGGCTGTGAGAGAACTCCAGACAGACACTGAA 2259
Oy      859 AspGlnGlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuHisSerAspPhe 878
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      2260 GAGACCTCCGAGACTTGTGAGAGAGATCGAGATCCTGAAATCTTGCAGCATGACAAC 2319
Oy      879 ILeValLysTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuArgLeuValMet 898
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      2320 ATCGTCAAGTCAAGAGGAGCTGCTACAGTGGGGTGGCGCAACCTTAAGATTATATG 2379
Oy      899 GluTyrLeuProSerGlyCysLeuArgAspLeuGlnArgHisArg---GlyLeuHis 917
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      2380 GAATATTTACATATGCAAGCTTTCAGAGACTATCTCCAAAACATAAAGACGATGAT 2439
Oy      918 ThrAspArgLeuLeuPheAlaTrpGlnIleCysGlyMetGluTyrLeuGlyAla 937
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      2440 CACAAAAAATCTTCAATACACATCTCAGATATGCAAGGAGGATGAAATATCTTGATACA 2499
Oy      938 ArgArgCysValHisArgAspLeuAlaIleArgAsnIleLeuValGluSerGluAlaHis 957
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      2500 AAAAGGTATATCCACAGGAGACTGSCACAAAGAAATATGTTGTAAGAGACAGC 2559
Oy      958 ValLysIleAlaAspPheGlyLeuAlaLysLeuLeuProleuGlyLysAspTyrTyrVal 977
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      2560 GTTAAATATGAGACTTGAGATTAAACAAGCTTGGCAGACAAAGAAATACTACAAA 2619
Oy      978 ValArgGluProGlyGlnSerProIlePheTrpTyrAlaProGluSerLeuSerAspAsn 997
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      2620 GTPAAGGAGCCAGGGGAAAGCCCATATTTCTGTGACGACCTGGAATCTTGCAGGAGC 2679
Oy      998 IlePheSerArgGlnSerAspValTrpSerPheGlyValValLeuTyrGluLeuPheThr 1017
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      2680 AAGTTTCTGTGGCCCAAGATGTGTGAGACTTGTGAGTGTCTATACGAACTTTTTCACA 2739
Oy      1018 TyrCysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetGlyProGluArg 1037
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      2740 TACATCGAGAGAGTAAAGTCCACCCGTGGAATATATGCGAATGATGATAA 2799
Oy      1038 GluGlyProProleuGlyS---ArgLeuLeuGluLeuAlaGluGlyArgArgLeuPro 1056
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      2800 CAAGGGCAATGATGTGTTCATTTGATAGAGCTTACTGAGAGCAACCGAAGATGGCA 2859
Oy      1057 ProProThrCysProThrGluValGlnGluLeuMetGlnLeuGlyStrAlaProGlu 1076
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      2860 AGCGCAAGAGATGCCAGATGAGATTTATGTATCATCAGACAGAGTGCAGAACACAT 2919
Oy      1077 ProHisAspArgProAlaPheAlaPheAlaThrLeuSer 1087
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      2920 GTGAGCCAGGCTCCCTTCAGAGACCTTCC 2952

```

```

RESULT 14
US-08-064-067D-2
: Sequence 2, Application US/08064067D
: Patent No. 5852184
: GENERAL INFORMATION:
: APPLICANT: Wilks, Andrew F.; Ziemlecki, Andrew;
: APPLICANT: Harpur, Ailsa
: TITLE OF INVENTION: No. 5852184el Protein Tyrosine Kinase
: NUMBER OF SEQUENCES: 23
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Felfe & Lynch
: STREET: 805 Third Avenue
: CITY: New York City
: STATE: New York
: COUNTRY: USA

ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/064,067D
FILING DATE: 30-Jun-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US91/08869
FILING DATE: 26-No. 5852184-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: Australian PK3594/90
FILING DATE: 28-No. 5852184-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: Australian 88229/91
FILING DATE: 27-No. 5852184-1991
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 5852184man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 5244
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-688-9200
TELEFAX: 212-638-3884
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 3495 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: nucleic acid
US-08-064-067D-2

Alignment Scores:
Pred. No.: 9,43e-210 Length: 3495
Score: 2360.00 Matches: 489
Percent Similarity: 65.89% Conservative: 164
Best Local Similarity: 49,348 Mismatches: 300
Query Match: 40,278 Gaps: 38
DB: 2 Indels: 15

US-09-397-967-16 (1-1099) x US-08-064-067D-2 (1-3495)
Oy      128 ILeuAspLeuHisValLeuGlnHisLeuPheAlaGlnHisArgSerAspLeuValSer 147
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      1 CTGCTGATGACTTTTCAATGCTCTTACCTTCCCGCAGAGCGGCATGATTTTGTTCAC 60
Oy      148 GlyArgLeuProValGlyLeuSerMetLysGluGlnGlyGluPheLeuSerLeuAlaVal 167
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      61 GGATGATTAAGATTAAGTCTGTGACTCATGAATCTCAGAAAGATGTGTTGGATGGCGTG 120
Oy      168 LeuAspLeuAlaGlnMetAlaArgGluGlnAlaGlnArgProGlyGluLeuLeuLysThr 187
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      121 TTAGCATGATGAGAAATAGCTTAAGAGAAACACAGACACTCCACTGCGTGTCTTAACCTG 180
Oy      188 ValSerTyrLysAlaCysLeuProProSerLeuArgAspValIleGlnGlyGlnAsnPhe 207
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      181 CTCAGCTCAAGACACTTCTTACCAAGTGCCTTGAGCGAAGATCCCAAGCTATTCACATT 240
Oy      208 ValThrArgArgArgIleArgArgThyValValLeuAlaLeuLeuProCysGlyArgLeu 227
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      241 TTAAACCGGAGAGCAATAGAGTACAGATTTGCGAGATTCATTCAGCAATTCAGTCAATGT 300
Oy      228 ProGlyArgProTyrAlaLeuMetAlaLysTyrIleLeuAspLeuGluArgLeuHisPro 247
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      301 AAAGCACTCCAGCAAGACCTAAAGTTAAGTATCTTAAACCTGGAACCTGCAGCTCT 360
Oy      248 AlaAlaThrThrGluThrPheArgVal-----GlyLeuProGlyAlaGln 262
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      361 GCCTTTCACAGACAGAGTTTGAAGTAAAGATCTGCAAGAGGT--CTTCAAGGTGAG 417
Oy      263 GluGluProGlyLeuLeuArgValAlaGlyAspAsnGlyIleProTrpSerSer----- 280

```


[illegible]

Qy	394	Thrrytllleuvaldagserserproglinsrtyrlyaserserphenleuthralacysval	413
Db	835	ctgatattactgcattgcattaccctccctaaggacttcacaaatattcttcgactccgttccgctt	894
Qy	414	glnthrproleugluproaerytyrlysglycysleuillaarglnasprosergilyla	433
Db	895	gagcgagaaaattgattattgaaattttaaacsactgttttcattacaaacgaatgacatggagag	954
Qy	434	phenSerleuvalglyleuserglnprohinsargSerLeuargglyleuvalaalaCys	453
Db	955	tacaaactcagtgaggacttaagacaaacttcagatgactttaaagacacttttgaaattgctpac	1014
Qy	454	trpanserglyleuvalasrvalalalaleuvaltytleuthsercysuylasrval	473
Db	1015	cagatggaaactggccctagacactatcattcttccacttaccaaatagctgcttcga	1074
Qy	474	argprolysglulysSerasphenleuvalvalargargglycysasnprohlaProla	493
Db	1075	aagccgaagaatgaattgaacttcatttcgttcctcaga-----acaaatggctgtttctgat	1128
Qy	494	proglycysSerprosercysCysala-----Leuthglnleuserphenisthr	510
Db	1129	gmpcagctctccaccacacatttaccagaccgcatatgaatgtaaaacaaatggctgttcacaa	1188
Qy	511	ileprothrapserserleuglutrphnlsyluasleuglnlsiglyserphenistrlysle	530
Db	1189	atcagaaatgacaaatttgatatttgaatgaagaacttgccscaagcacactttttacaaaatga	1248
Qy	531	phenarglyseratargatrglvalalaser--glyluthninsaspserglvalleu	549
Db	1249	ttttaaagcctgaagaaacagaaacttggagactttagtgcactccacgaaacccgaactgttt	1308
Qy	550	leuylsvalmetaspserserlagnhsatgandancysmetglnserphenleuglnalalasr	569
Db	1309	tttgaaagctccagatgaaacagacacatpbaaactatgttcagactgtttgttgaaacacaaagc	1368
Qy	570	leuuserglnvalsertyrprohinsleuvalleuileuinsglylvalcysmetalaagly	589
Db	1369	atgatgatgactgactttctscacaaagcatttggcttgaattatgaaatgatgtctgtcgga	1428
Qy	590	---aspsertleuethvaldnglnluphrenaltyrleuglualalalasrmetelyleuarg	608
Db	1429	cagagaaacatttttggttccaaaggatttgaaaaattttggatccctgaaatccctgaag	1488
Qy	609	lysargllynlisleuvalserlasertrpilyserleuglnvalthrysglnleuvaltyr	628
Db	1489	aagaaacaaaatttctataaattatattatgtgaaacttggaagtcggcgaacacacttgccactgc	1548
Qy	629	AlaleuasnTyrleuglnasrplysglyleuprohinsglyAsnvalseralarglyval	648
Db	1549	gccattgcacttccctccgaaagaaaatccctttattcattcagcgaaatgctgtgcctaaatattgc	1608
Qy	649	leuileuvalargluclylgluyl-----Asrglyasnprohophleileysleuserasp	666
Db	1609	ctgcctttatccagaaagaaacagacagaaacagcgagaaacccacttttcattcaaaacttatgat	1668
Qy	667	proglyvalserProthrylValleu-----Serleuglnmetleu	679
Db	1669	cctggcacttggacatttaccagtttaccgaagaaactttctctcgttggtttccaaactgttt	1728
Qy	680	thrsaprrarglterprotrpvalalalaproglucysleuglnlualaglnthrlencusyleu	699
Db	1729	cagagaaagaaatccattgggtaacacactcgaagcattggaatccttAAAATCTAACTCTG	1788
Qy	700	gluilaasplystrpcllypheniglyalatatrthtrpglivalalpheniglnargglyProla	719
Db	1789	gcAACAGACAAGtgagacttgcggagacacactcgttgagaaatgtgcgaatggagaaataga	1848
Qy	720	hislterthrsleugluproalalalyslysleuylsphenetyrGluaspnglnclylneu	739
Db	1849	ccccctgactgctctggcgaatttctcaaaaagactctgcacttttaagaagatgaactcagctt	1908

QY	740	ProAlaLeuLysTrpPthrGlnLeuAlaGlyLeuIleThrGlnCysMetAlaTrpAspPro	759
Db	1909	CGTGACCCAAAGTGGACAGAGTTGGCCAAACCTTTAAATTAATTTGCATGACATGACGACA	1968
QY	760	GlyArGArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIleThrSerAsp	779
Db	1969	GATTTCAGGCGCTGCTTTCACAGCTGTTCATCCGGATCTTAAACAGGCTGTTTACTCCAGAT	2028
QY	780	TyrGlnLeuLeuSerAspProThrProGlyIleProSerProArg--AspGlnLeuCys	798
Db	2029	TATGACACTACTAAACAGAA---AATGACATGCTCCACAAACATGACATGATGCTCCCTACGG	2085
QY	799	ValAlaGlyAlaGlnLeuLysTrpAlaCysGlnAspProAlaIlePheGlnGlnAArgHisLeu	818
Db	2086	TTTTCTGTGCTGCT-----TTTGAAGACAGGGACCCCTACACAGTTGAAGAGACACTTG	2139
QY	819	LysTrpLileSerLeuLeuGlyLysGlyAsnPheGlySerValGlnLeuCysArGTrpAsp	838
Db	2140	AAGTTTCAAGCAGCTTGTGGCAAAAGTAACCTTGCGGAGTGTGAGATGGCCGCTATAC	2199
QY	839	ProLeuGlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGlnHisSerAlaPro	858
Db	2200	CCGTCGACGAGCAACACTGGCGAGGTGGTCGCTGTAAACAACTCCAGCACAGCACTGAA	2259
QY	859	AspGlnGlnAspPheGlnArgGlnIleGlnIleLeuLysAlaLeuHisSerAspPhe	878
Db	2260	GAGCACCTCCGACAGCTTTGAGAGGAGATCGAGATCTCTAATTCCTGACATGACAC	2319
QY	879	IleValLysTrpArgGlyValSerTrpGlyProGlyArgGlnSerLeuArgLeuValMet	898
Db	2320	ATCGTCAAGTCAAGGGAGGTGCTCTACAGTGGCGGGCGCCGACCTTAAGATTATTTATG	2379
QY	899	GluTrpLeuProSerGlyCysLeuArgAspLeuLeuGlnArgHisArg--GlyLeuHis	917
Db	2380	GAAATTATTACCATATGGAGACTTACGAGATATCTCAAAACATTAAGAACGGTTAAT	2439
QY	918	ThrAspArgLeuLeuLeuPheAlaTrpGlnIleCysLysGlyMetGlnTrpLeuGlyAla	937
Db	2440	CACAAAAAAGCTTTTCATATACACATCTCAGATATGCAGAGCATGAAATATCTTGTCACA	2499
QY	938	ArgArgCysValHisArgAspLeuAlaIleArgAsnIleLeuValGlnSerGlnAlaHis	957
Db	2500	AAAAGGATATTCGCCAGGGAGCTGGCAACAGGAACATTTTGGTGCAAAATGCAAGACAG	2559
QY	958	ValLysIleAlaAspPheGlyLeuAlaIleLysLeuLeuProLeuGlyLysAspTrpTrpAl	977
Db	2560	GTTAAATAGAGACGACTTCGGATTACCAAAAGTCTTCCCGACAGACAAAGAAATCTACAA	2619
QY	978	ValArgGlnProGlyGlnSerProIlePheTrpTrpAlaProGlnSerLeuSerAspAsn	997
Db	2620	GTAAGGCGCCAGGGAGAAACCCCAATATCTGTGACGCACTGAAATCTTCGACGAGAGC	2679
QY	998	IlePheSerArgGlnSerAspValIleTrpSerPheGlyValValLeuLysGlnLeuPheThr	1017
Db	2660	AAGTTTCTGTGGCCTCAGATGTGTGGAGGTTTGGAGTGGTCTTATACCAACTTTTCACA	2739
QY	1018	TyrCysAspLysSerCysSerProSerAlaGlnPheLeuArgMetMetCylProGlnArg	1037
Db	2740	TATCATCGAGAAAGATTAAAGATCCACCCGCGTAATTTATCGAATGATTGGCAATCATANA	2799
QY	1038	GlnGlyProProLeuCys---ArgLeuLeuGlnLeuLeuAlaGlnGlyArgArgLeuPro	1056
Db	2800	CAGAAGCAAAATGATTGTTCATTTGTATAGACATCTAAGACGACAGGAAGATTTCGA	2859
QY	1057	ProProProThrCysProThrGlnValaGlnGlnLeuMetGlnLeuCysTrpAlaProGln	1076
Db	2860	AGGCCAAGAAGATGCCCAAGATAGAGATTTAATGTATCATATACAGAGACTGCGTAACAAAT	2919
QY	1077	ProHisAspArgProAlaPheAlaThrLeuSer	1087
Db	2920	GTGAGCCAGCGCTCCCTCTCCAGGAGCACTTTTCC	2952

Tue Apr 29 06:04:28 2003

us-09-397-967-16.rni

Page 30

Search completed: April 28, 2003, 17:00:17
Job time : 296.669 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 28, 2003, 18:32:45 ; Search time 3018 Seconds

(without alignments)
5897.561 Million cell updates/sec

Title: US-09-397-967-16

Perfect score: 1099

Sequence: 1 MAPSEPTPLIPORCSLSS.....RPAFATISPOLDPLMRGPG 1099

Scoring table:

OLIGO	
Xgapop 60.0 , Xgapext 60.0	
Ygapop 60.0 , Ygapext 60.0	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 16154066 seqs, 8097743376 residues

Word size: 1

Total number of hits satisfying chosen parameters: 32308016

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: listing first 150 summaries

Command line parameters:

```

-MODEL=frame_plus_p2n_model -DEV=xlp
-O=/gen2.1/USPRO.spool/US09397967/runat_26042003.182927.11215/app.query.fasta.1.1287
-DB-RST -QEXT=fastap -SUFFIX=oligo_rst -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=150
-DOCLIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=100 -MODE=LOCAL -OUTPMT=ptc
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09397967 @GNN.1.1.2605 @runat_26042003.182927.11215 -NCP=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV.TIMEOUT=10 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

```

Database :

```

ESR:*
1: em_estba:*
2: em_esthum:*
3: em_estlm:*
4: em_estlmv:*
5: em_estlov:*
6: em_estlpl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	152	13.8	971	13 B1411962	B1411962 602966375
2	149	13.6	580	10 BB608771	BB608771 BB608771
3	142	12.9	583	10 BB631706	BB631706 BB631706
4	142	12.9	720	12 BG873355	BG873355 602794328
5	135	12.3	1625	11 BC027234	BC027234 Mus muscu
6	110	10.0	512	9 AA510093	AA510093 vg3sc07.r
7	92	8.4	545	12 BE916148	BE916148 601666763
8	88	8.0	826	13 B157690	B157690 603237050
9	82	7.5	675	12 BF150250	BF150250 uy8e09.y
10	77	7.0	907	12 BG870513	BG870513 602791432
11	74	6.7	453	9 AA023709	AA023709 mh7e02.r
12	74	6.7	478	9 AA023670	AA023670 mh78906.r
13	73	6.6	593	9 AA881653	AA881653 vx21c02.r
14	71	6.5	698	14 BM935182	BM935182 UT-M-BH3-
15	69	6.3	436	12 BF181053	BF181053 601805105
16	65	5.9	436	10 AW141115	AW141115 EST291147
17	65	5.9	590	14 W48204	W48204 mc87h11.r1
18	62	5.6	559	12 BG800853	BG800853 0053-91.M
19	60	5.5	358	13 B153609	B153609 602871178
20	56	5.1	621	10 AW742526	AW742526 ups8e12.y
21	53	4.8	539	12 BE913247	BE913247 601668246
22	51	4.6	867	9 AA755769	AA755769 vv35a11.r
23	49	4.5	543	12 BE753801	BE753801 206963.MA
24	49	4.5	557	13 B1339676	B1339676 364724.MA
25	46	4.2	544	12 BF705690	BF705690 243606.MA
26	46	4.2	562	12 BF193178	BF193178 244670.MA
27	45	4.1	659	17 AE936543	AE936543 2M0223004
28	44	4.0	417	13 B1540725	B1540725 454105.MA
29	44	4.0	441	9 AA020387	AA020387 mh60c03.r
30	44	4.0	501	10 BE623895	BE623895 ut14d09.y
31	44	4.0	597	9 AA015273	AA015273 mh30h10.r
32	42	3.8	581	12 BF149581	BF149581 uy66f08.y
33	41	3.7	554	12 BF191625	BF191625 239261.MA
34	39	3.5	731	9 AA920299	AA920299 vx93g02.r
35	38	3.5	410	9 AA023069	AA023069 mh66d02.r
36	35	3.2	308	9 AA174246	AA174246 mo95a10.r
37	34	3.1	231	10 AW501263	AW501263 UT-HF-BPO
38	34	3.1	254	10 AW501582	AW501582 UT-HF-BPO
39	34	3.1	286	10 AW502902	AW502902 UT-HF-BPO
40	34	3.1	437	10 AW501452	AW501452 UT-HF-BPO
41	34	3.1	451	10 AW500984	AW500984 UT-HF-BPO
42	34	3.1	462	10 AW408832	AW408832 UT-HF-BPO
43	31	2.8	478	12 BF546077	BF546077 UT-R-C2P-
44	30	2.7	483	10 AW462965	AW462965 BP230011A
45	30	2.7	763	9 AJ399251	AJ399251 AJ399251
46	30	2.7	769	9 AJ442008	AJ442008 AJ442008
47	28	2.5	411	10 AW503697	AW503697 UT-HF-BNO
48	27	2.5	491	12 BF441858	BF441858 258301.MA
49	27	2.5	550	10 BE309453	BE309453 601095351
50	25	2.3	526	12 BF191620	BF191620 239255.MA
51	25	2.3	564	13 BM088287	BM088287 501721.MA
52	25	2.3	871	14 BQ709825	BQ709825 AGENCOURT
53	24	2.2	371	10 BB791980	BB791980 BB791980
54	23	2.1	339	10 AW230832	AW230832 uc68d02.y
55	22	2.0	540	10 AW505513	AW505513 UT-HF-BNO
56	21	1.9	472	10 AW785501	AW785501 116874.MA
57	18	1.6	258	10 BB590115	BB590115 BB590115
58	18	1.6	291	9 A1651109	A1651109 wa97d05.x
59	18	1.6	319	13 BM032194	BM032194 497815.MA
60	18	1.6	373	13 BM146700	BM146700 TCAAP1E70
61	18	1.6	413	10 AW504672	AW504672 UT-HF-BNO
62	18	1.6	481	17 AO192794	AO192794 HS-2251.B
63	18	1.6	781	13 B1906136	B1906136 603063556
64	18	1.6	814	13 B1911448	B1911448 603063362

```

C 65 17 1.5 313 12 BF147824 uY83609.x
C 66 17 1.5 426 13 BG995133 CM0-HT129
C 67 17 1.5 430 9 A1472755 tA13e03.x
C 68 17 1.5 449 9 A1596913 EST250616
C 69 16 1.5 239 10 BB285471 BB285471
C 70 16 1.5 258 12 BF367544 BF367544
C 71 16 1.5 304 10 BB285470 BB285470
C 72 16 1.5 398 9 AA453345 AA453345
C 73 16 1.5 406 14 T48592 Ph6b6_19/1T
C 74 16 1.5 430 13 BM147781 BM147781
C 75 16 1.5 532 12 BG364706 BG364706
C 76 16 1.5 612 9 AA521703 AA521703
C 77 16 1.5 753 12 BG870416 BG870416
C 78 16 1.5 766 13 B1153342 B1153342
C 79 16 1.5 920 14 BO898982 BO898982
C 80 16 1.5 945 14 BO732700 BO732700
C 81 16 1.5 989 17 CNS042XO CNS042XO
C 82 16 1.5 1001 17 CNS038MW CNS038MW
C 83 15 1.4 253 12 BF373713 BF373713
C 84 15 1.4 754 17 BH043390 BH043390
C 85 14 1.3 281 10 AM437524 AM437524
C 86 14 1.3 296 10 AM212845 AM212845
C 87 14 1.3 318 9 AA008555 AA008555
C 88 14 1.3 449 10 BF711016 BF711016
C 89 14 1.3 636 14 BO780982 BO780982
C 90 14 1.3 683 14 BO781477 BO781477
C 91 13 1.2 381 13 B1643881 B1643881
C 92 13 1.2 441 12 BG509232 BG509232
C 93 13 1.2 444 14 BM813523 BM813523
C 94 13 1.2 500 13 B1424599 B1424599
C 95 13 1.2 562 10 AM21037 uO70a12.y
C 96 12 1.1 145 10 BE156215 BE156215
C 97 12 1.1 214 9 AA320481 AA320481
C 98 12 1.1 270 14 H24996 H24996
C 99 12 1.1 294 13 B1051048 B1051048
C 100 12 1.1 308 10 BB605196 BB605196
C 101 12 1.1 316 10 BE554456 BE554456
C 102 12 1.1 318 12 BF741156 BF741156
C 103 12 1.1 324 14 D77965 D77965
C 104 12 1.1 336 9 AA069054 AA069054
C 105 12 1.1 346 12 BF886669 BF886669
C 106 12 1.1 353 9 AA715499 AA715499
C 107 12 1.1 354 17 BH826047 BH826047
C 108 12 1.1 355 10 AM059494 AM059494
C 109 12 1.1 366 9 AA233516 AA233516
C 110 12 1.1 372 9 AA572234 AA572234
C 111 12 1.1 383 9 A1138105 A1138105
C 112 12 1.1 392 9 A1483943 A1483943
C 113 12 1.1 399 14 R17980 R17980
C 114 12 1.1 409 17 CNS040CZ CNS040CZ
C 115 12 1.1 437 14 BM823021 BM823021
C 116 12 1.1 437 14 BM857398 BM857398
C 117 12 1.1 438 10 AV978849 AV978849
C 118 12 1.1 445 13 BG987835 BG987835
C 119 12 1.1 453 12 BF391056 BF391056
C 120 12 1.1 462 13 B1221056 B1221056
C 121 12 1.1 462 13 BM288735 BM288735
C 122 12 1.1 472 12 BF774038 BF774038
C 123 12 1.1 473 12 BG458915 BG458915
C 124 12 1.1 475 14 BM968652 BM968652
C 125 12 1.1 479 12 BF769324 BF769324
C 126 12 1.1 488 9 A1444260 A1444260
C 127 12 1.1 496 10 AM665954 AM665954
C 128 12 1.1 497 14 BO560624 BO560624
C 129 12 1.1 499 10 AM469824 AM469824
C 130 12 1.1 499 17 A2493659 A2493659
C 131 12 1.1 502 10 AM632498 AM632498
C 132 12 1.1 504 13 BJ478073 BJ478073
C 133 12 1.1 504 14 R56713 R56713
C 134 12 1.1 516 9 A1150337 A1150337
C 135 12 1.1 517 10 AM938783 AM938783
C 136 12 1.1 518 14 BM756676 BM756676
C 137 12 1.1 520 12 BE872391 BE872391

```

```

138 12 1.1 523 10 AM489577 UI-M-BH3-
C 139 12 1.1 526 12 BF736600 BF736600
C 140 12 1.1 528 17 CNS01XHO A1171693 Tetradon
C 141 12 1.1 532 12 BG798512 BG798512
C 142 12 1.1 532 17 AO377376 AO377376
C 143 12 1.1 536 14 BM756648 BM756648
C 144 12 1.1 538 12 BG812095 BG812095
C 145 12 1.1 543 10 BE330992 BE330992
C 146 12 1.1 546 12 BF722796 BF722796
C 147 12 1.1 546 14 BM742052 BM742052
C 148 12 1.1 547 14 BQ399048 BQ399048
C 149 12 1.1 548 14 BM756028 BM756028
C 150 12 1.1 549 10 BE632486 BE632486

ALIGNMENTS

RESULT 1
BI411962 971 bp mRNA linear EST 14-AUG-2001
LOCUS BI411962
DEFINITION 602966375F1 NCI_CGAP_Lu33 Mus musculus cDNA clone IMAGE:5121786 5',
mRNA sequence.
ACCESSION BI411962
VERSION BI411962.1 GI:15172885
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryotes; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 971)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: gcaps-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: L1AM1297 row: e column: 19
High quality sequence start: 25
High quality sequence stop: 785.
Location/Qualifiers
1. 971
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:5121786"
/clone_lib="NCI_CGAP_Lu33"
/tissue_type="pooled lung tumors"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pT73D-Pac (Pharmacia) with a
modified polylinker; site_1: NotI; site_2: EcoRI; 1st
strand cDNA was prepared from mRNA obtained from pooled
lung tumors with a Not I - oligo(dT) primer [5'
TGTTCACATCTCAAGTGGAGCGCCGCTCTGTTTTTTTTTTT 3']
Double-stranded cDNA was ligated to Eco RI-adaptors
(Pharmacia), digested with Not I and cloned into the Not
I and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT 216 a 274 c 279 g 202 t
ORIGIN
Alignment Scores: 1.19e-137 length: 971
Score: 152.00 Matches: 196
Percent Similarity: 98.99% Conservative: 0
Best Local Similarity: 98.99% Mismatches: 1

```

Query Match:	13.83%	Indels:	2
DB:	13	Gaps:	0
US-09-397-967-16 (1-1099) x B1411962 (1-971)			
QY 490	ProAlaProAlaProG1CYSerProSerCySCysAlaLeuThrGlnLeuSerPheHis	509	
Db 51	CCCCGGCCTGCCCTGGCTGCTCCCCCTCTCTCTGCGCTACACACAGTGAAGTCCAC	110	
QY 510	ThrIleProThrAspSerLeuGlnThrPheAsnLeuGlnAspGlySerPheThrIys	529	
Db 111	ACAATTCCACAGCAGCAGCTGGATGGACGACGAGAACCTGGGTCACGGTTCTTTACCAAG	170	
QY 530	IlePheAlaGlySer-ArgArgGluValValAspGlyGlnThrHisAspSerGluValle	549	
Db 171	ATCTTCCTGGCCG-CAGGCGGAGAGTGGTGGATGGTGAACACATGACTCGGAAGTCTT	229	
QY 549	UleuLysValIleCAspSerArgHisArgAsnGlyMetGluSerPheGluAlaAlaIse	569	
Db 220	CCTGAAGCTCATGAGCTCTCAGACATCGGAACTGCATGGAGTCTTTTCTGGAAGCCGCAAG	289	
QY 569	rLeuMetSerGlnValSerTyrProHisLeuValLeuLeuHisGlyValAlcysMetAlaG1	589	
Db 290	CTTGATGAGCCACAGATACCTACCCGACCTGGTGTACTAGCAGCGGCGTGCATGGCTGG	349	
QY 589	YAspSerIleMetValGlnGlnPheValItyrLeuGlnValAlaIleAspMetTyrLeuArgly	609	
Db 350	AGACGACATCATGGGACGAGAAATTGTGTATCTAGGACCAATGTGACATGTACCTCGCA	409	
QY 609	sArgGlyHisLeuValSerAlaSerTyrPheGlnValThrIysGlnLeuAlaItyrAl	629	
Db 410	GGGTGGCCACCTCGGTGTACGCCAGCTGGAAATCGAGTGAACCAAGCACTGGCATATGC	469	
QY 629	aLeuAsnTyrLeuGlnAspGlyLeuPheProHisGlyAsnValSerAlaArgLysValle	649	
Db 470	CCTTACTACTGGAGGACAAAGGCTTCTCAGCGCAACGCTCAGCAGGAAAGTGT	529	
QY 649	UleuAlaArgGlnGlyGlyAspGlyAsnProPheIleLeuLysLeuSerAspProGlyIya	669	
Db 530	CTGCTCTGTGAGGGGGGTGATGGAAATCCACCTTTCATTAAGCGAGTGCCTGGTGT	589	
QY 669	IserProThrValLeuSerLeuGlnMetLeuThrAspArgIleProTyrPval	686	
Db 590	CAGTCCCATCTGCTGAGCCTGGAAATGCTCACGCCACAGATACCTGGGTG	641	
RESULT 2			
LOCUS	BB608771	580 bp	mRNA linear EST 26-OCT-2001
DEFINITION	BB608771 RIKEN full-length enriched, 2 days pregnant adult female		
ACCESSION	BB608771		
VERSION	BB608771.2	GI:16451108	
SOURCE	house mouse.		
ORGANISM	Mus musculus		
REFERENCE	1 (bases 1 to 580)		
AUTHORS	Arkhamov, K., Carlini, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Kouda, M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K., Ohno, M., Okazaki, Y., Okido, T., Salto, R., Sakai, C., Sakai, K., Sano, H., Sasakhi, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, T., Toya, T., Muramatsu, M., and Hayashizaki, Y.		
TITLE	RIKEN Mouse ESTs (Arakawa, T., et al. 2001)		
JOURNAL	Unpublished (2001)		
COMMENT	On Dec 6, 2000 this sequence version replaced gi:11563948. Contact: Yoshihide Hayashizaki Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute The Institute of Physical and Chemical Research (RIKEN) 1-7-22 Stenhiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan		

US-09-397-967-16 (1-1099) x BB608771 (1-580)

1 Meta[apProSe]glUgltHrPrObLeuIepProGInArGerCySerIeUsErSer 20

US-09-397-967-16 (1-1099) x BB608771 (1-580)

1 Meta[apProSe]glUgltHrPrObLeuIepProGInArGerCySerIeUsErSer 20

Db	73	ATGGACCTCCAGAGGAGAGACACCTCTGATCCCTCACCGCTCTTGACGCTTCATCC	132
Qy	21	SerGluAlaGlyAlaLeuHisValLeuLeuProProArgGlyProGlyProProGlnArg	40
Db	133	TCAGAGGACGAGGACCCCTGCATGTGCTCTCTCCGCCGGAGCATGGGCTCCACAGCA	192
Qy	41	LeuSerPheSerPheGlyAspTyrIleuAlaGluAspLeuGlyValArgAlaAlaLysAla	60
Db	193	TTGTCAATCTCTTTTGGGGAGACTTGGCTGGAGATTATGTGTCCGAGCTCCAAAGGC	252
Qy	61	CysGlyIleLeuProValTyrHisSerIleuPheAlaLeuAlaThrGluAspPheSerCys	80
Db	253	TGTGGCATCTCCCGTGGTTATTCATTCGGCTTTCGGCTCCGACGAGGACTTCCTTTC	312
Qy	81	TrpPheProProSerHisIlePheCysIleGluAspValAspThrGlnValLeuValTyr	100
Db	313	TGGTTCCTCCCAAGCCACACATCTTCTGCATAGAGAGCGTGGACATCACTTGGTCTAC	372
Qy	101	ArgLeuArgPheTyrPheProAspTrpPheGlyLeuGluThrCysHisArgPheGlyLeu	120
Db	373	AGGCTACGCTTTATTTCCTGCATCGTGTGGGCTGGAGACATGTCCACCGCTTTGGCGG	432
Qy	121	ArgLysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLeuPheAlaGln	140
Db	433	CGCAAGATTGTACACAGTGCATCTTGACTTACACGTTTGTAGAACATCTTTGCTCAG	492
Qy	141	HisArgSerAspLeuValSerGlyArg	149
Db	493	CACCGCAGTGCCTGTGTAGTGGGGC	519

RESULT 3							
BB631706							
LOCUS	BB631706	583 bp	mRNA	linear	EST 26-OCT-2001		
DEFINITION	BB631706	RIKEN full-length cDNA clone A130091E14 5',	16 days neonate	thymus	Mus musculus		
		CDNA clone A130091E14 5',	mRNA sequence.				

Accession	BB631706	
Version	BB631706.1	GI:16468415
Keywords	EST.	
Source	house mouse.	
Organism	Mus musculus	

REFERENCE
1 (bases 1 to 583)
Arakawa, T., Carinci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A.

TITLE
RIKEN Mouse ESTs (Arikawa, T., et al. 2001)

TITLE	RIKEN Mouse ESTs (Arakawa, T., et al. 2001)
JOURNAL	Unpublished (2001)
COMMENT	Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel.: 01-45-503-0222

Tel: 01-45-503-9222
 Fax: 01-45-503-9216
 Email: genome-rs@ncs.riken.go.jp

URL:<http://genome.gsc.riken.go.jp/>
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh

M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new genes. *Genome Res.* 10 (10), 1617-1630 (2000)

wagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura

S., Kawai, J., Okazaki, I., Mutsaers, M., Inoue, I., Nita, A. and Hayashizaki, Y.
Open integrated sequence analysis (OISA) custom--384-format

- NINE integrated sequence analysis (NISA) system 304 Local sequencing pipeline with 384 multiplexed sequencer. Genome Res.

10 (11), 1757-1771 (2000)
Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara, Y. and Hayashizaki, Y.
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. *Genome Res.* 11 (2), 281-289 (2001)
Kondo, S., Shingawa, A., Saito, T., Kiyosawa, H., Yamamoto, I., Aizawa, K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and Hayashizaki, Y.
Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences. *Mamm. Genome* 12, 673-677 (2001)
Please visit our web site (<http://genome.9sc.riken.go.jp>) for further details.
e mouse tissues.

```

FEATURES
SOURCE
Location/Variables
1. .583
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="A130091E14"
/clone.lib="RIKEN full-length enriched, 16 days neonate
thymus"
/tissue_type="thymus"
/dev_stage="16 days neonate"
/lab_host="DH10B"
/note="Site.1: SalI; Site.2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was

```

GAGAGAGAGAAGGATCCAGAGGCTTTTTTTTTTTTTTTTTV N 3'), cDNA was prepared by using triethose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 185.0. Second strand cDNA was prepared with the primer adapter of sequence (5' GAGAGAGAGATTCGAGTATATTAATATACCCCCCCCC 3'). cDNA was cleaved with XhoI and BamHI. Vector: a modified pBluescript KS(+) after bulk excision from lambda FLC I."

BASE COUNT	106	a	171	c	156	g	150	t
ORIGIN								
Alignment Scores:								
Pred. No.:	4,01e-128	Length:	583					
Score:	142.00	Matches:	142					
Percent Similarity:	100.00%	Conservative:	0					
Best Local Similarity:	100.00%	Mismatches:	0					
Query Match:	12.92%	Indels:	0					
DB:	10	Gaps:	0					

US-09-397-967-16 (1-1099) x BB631706 (1-583)

QY 1 MetaAlProPseSerGluGlnrPrroLeuIleProGlnArgSerCysSerLeuSerSer 20
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
60 MGGAGTGTCTGATGTAAGCGAAGACGTCGACCCGCCGCGGCAACCCGCTCATGCC 81

```

21  Soc[na]a[va]a[un]i[s]a[le]u[Pro]bra[ng]i[bro]gi[bro]gi[na] 40
DB  22  ATGGCACCCTCCAGTGGAGAGACCCCTGATCTCCTCAGGGCTTCAGGCCCTCCATCC  81

```

82 TCAGAGGAGAGCCCTGCATGTGCTCTTCTCCGCCGGGAGACTGGGCTCCCCACGGG 141

41 LeuSerPheSerPheGlyAspTyrLeuValGluAspLeuCysValArgAlaIaIaLysAla 60

Db 142 TTGTCATTCTTTTGGGACTTGGCTGAGATTATGTGTGCGACTGCCAAGCC 201

61 CysGlyIleLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspSerCys 80

Db 202 TGTGGCAATCCTGCTGTTTATCATTCGCTTTTCGCTGCGCACTGAGGACITTCCTTGG 261

QY 81 TrpPheProSerHisIlePheCysTleGluAspValAspTrpGlnValLeuValTyr 100

Db 262 TGGTTTCCCAAGCCACATCTTCTGCATAGAGGACGCGACACTCAAGTCTTGGTCTAC 321

OY 101 ArgLeuArgPheThrPheProAspTrpPheGlyLeuGluThrcysHisArgPheGlyLeu 120
 DB 322 AGGCTACGCTTTTATTTTCCCTGACTGCTGTTGGGCTGAGACATGTCACCGCTTTGGGCTG 361
 OY 121 ArgGlyAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLeuPheAlaGln 140
 DB 382 CGCAAGATTTGACACAGTGCATCTTGACTACAGCTTTTACAGATCTCTTGGCTCAG 441
 OY 141 HisArg 142
 DB 442 CACCGC 447

RESULT 4
 BG873355 720 bp mRNA linear EST 29-MAY-2001
 LOCUS BG873355
 DEFINITION 602794328F1 NCI_CGAP_SG2 Mus musculus cDNA clone IMAGE:4925598 5',
 mRNA sequence.
 ACCESSION BG873355
 VERSION BG873355.1 GI:14223895
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

REFERENCE NIH-MGC http://mgc.nci.nih.gov/
 1 (bases 1 to 720)
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished (1999)
 JOURNAL Contact: Robert Strausberg, Ph.D.
 COMMENT Email: cgabbs-r@mail.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LAM10850 row: 9 column: 07
 High quality sequence stop: 709.

FEATURES
 source location/Qualifiers
 1..720

/organism="Mus musculus"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:4925598"
 /clone_lib="NCI_CGAP_SG2"
 /lab_host="DH10B (T1 phage-resistant)"
 /note="Organ: salivary gland; Vector: pCMV-Sport6; Site: 1;
 Note: Site 2: Salivary gland; Cloned unidirectionally. Primer: Oligo
 dT. Average insert size 1.3 kb. Constructed by Life
 Technologies. Note: this is a NCI_CGAP Library."
 BASE COUNT 123 a 239 c 229 g 129 t
 ORIGIN

Alignment Scores:
 Pred. No.: 5,15e-128 Length: 720
 Score: 142.00 Matches: 142
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 12.92% Indels: 0
 DB: 12 Gaps: 0

US-09-397-967-16 (1-1099) x BG873355 (1-720)

OY 916 LeuHisThrAspArgLeuLeuLeuPheAlaTrpGlnIleCysGlyGlyMetGluTrpLeu 935
 DB 79 CTGACACCGACCGCTACGCTGCTTCGTCGACATCTGCAAGGCGCATGAGTACCTG 138
 OY 936 GLYALARGARGCYValHisArgAspLeuAlaAlaArgAsnIleLeuValGluSerGlu 955
 DB 139 GGTGGCGCGCGCTGCTACACCGTGCCTGCGCGGCAACATCTTGGTGGAGAGCGAG 198

OY 956 AlaHisValIleAlaAspPheGlyLeuAlaIleLeuAspProLeuGlyAspTrp 975
 DB 199 GCTCATGTGAAGATCCGGACTTCGGCTTCGCTGACTGCTGCTGCGGAAAGCACTAC 258
 OY 976 TyrValValArgGluProGlyGlnSerProIlePheTrpTyrAlaProGluSerLeuSer 995
 DB 259 TACGTGCTCCGCGAGCTGCGCAAGCCCATCTTTGGATATGCCGCGAGTCCCTATCT 318
 OY 996 AspAsnIlePheSerArgGlnSerAspValTrpSerPheGlyValValLeuTrpGluLeu 1015
 DB 319 GACACATCTTCTCCGCGCAATCTGAGCTTGGAGCTTGGAGTGTGTTCAGAGCTC 378
 OY 1016 PheThrTyrCysAspIleSerCysSerProSerAlaGluPheLeuAlaGlyMetGlyPro 1035
 DB 379 TTCACCTACGCGACAAAGACTCTGAGCCCATCGCGTGTGTTCTCGGCATGATGGGCTT 438
 OY 1036 GluArgGluGlyProProLeuCysArgLeuGluGluLeuAlaGluGlyArgGluLeu 1055
 DB 439 GAGCGTGAAGAGCCCGCGCTGCTGAGCTGCTGAGCTGCTGAGCGCGAGCGCTC 498
 OY 1056 ProPro 1057
 DB 499 CCACCA 504

RESULT 5
 BC027234 1625 bp mRNA linear HTC 07-AUG-2002
 LOCUS BC027234
 DEFINITION BC027234
 ACCESSION BC027234
 VERSION BC027234.1 GI:20071020
 KEYWORDS HTC.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE NIH-MGC Project URL: http://mgc.nci.nih.gov
 1 (bases 1 to 1625)
 AUTHORS Strausberg, R.
 TITLE Direct Submission
 JOURNAL Submitted (04-APR-2002) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA

REMARK
 COMMENT NIH-MGC Project URL: http://mgc.nci.nih.gov
 Contact: MGC help desk
 Email: cgabbs-r@mail.nih.gov
 Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Baylor College of Medicine Human Genome
 Sequencing Center
 Center code: BCM-HGSC
 Web site: http://www.hgsc.bcm.tmc.edu/cdna/
 Contact: amgebcm.tmc.edu
 Gunaratne, P.H., Garcia, A.M., Lu, X., Huliyk, S.W., Hale, S.M.,
 Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Wuzny, D.M.,
 Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
 Series: IRK Plate: 35 Row: 9 Column: 19
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Similarity but not
 identity to protein
 This clone has the following problem: Incomplete processing.

FEATURES
 source location/Qualifiers
 1..1625

/organism="Mus musculus"
 /db_xref="taxon:10090"
 /map="C57BL/6J"
 /clone="IMAGE:3489805"
 /tissue_type="mammary tumor. WAP-Trf alpha model. 7 months
 old, gross tissue."
 /clone_lib="NCI_CGAP_Mam5"

/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"

BASE COUNT 328 a 470 c 483 g 344 t

ORIGIN

Alignment Scores:

Pred. No.:	Length:	Matches:
9.65e-121	1625	186
Score: 135.00		
Percent Similarity: 97.89%	Conservative: 0	
Best local Similarity: 97.89%	Mismatches: 2	
Query Match: 12.28%	Indels: 4	
DB: 11	Gaps: 0	

US-09-397-967-16 (1-1099) x BC027234 (1-1625)

QY 1 MetAlaProSerGluGluThrProLeuLeuProGlnArgSerCysSerLeuSer 20
 DB 42 ATGGACCTCCAGAGAGAGACACCTCTGATCCCTCAGCGCTTTCAGCCTTCATCC 101
 QY 21 SerGluAlaGlyAlaLeuHisValLeuLeuProProArgGlyProGlyProGlnArg 40
 DB 102 TCAGAGGAGAGAGCGCTCATGTGCTCTCCCTCCCGGGGACCTGGGCTCCCGACGGA 161
 QY 41 LeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysValArgAlaAlaLysAla 60
 DB 162 TTGCATATCTCTTTGGGAGCTACTTGGCTGAGATTATGTGTCGAGCTGCCAAGGCC 221
 QY 61 CysGlyLeuLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSerCys 80
 DB 222 TGTGGCACCCTGCTGTTATCATCTGCTTTGCTGCTGCGACCTGAGGACTTCTTTC 281
 QY 81 TrpPheProProSerHisIlePheCysIleGluAspValAspThrGlnValLeuValTyr 100
 DB 282 TGGTTTCCCAAGACCATCTTCTGATAGAGACGTGACACACTCAAGTCTTGCTTAC 341
 QY 101 ArgLeuArgPheTyrPheProAspTrpPheGlyLeuGluThrCysHisArgPheGlyLeu 120
 DB 342 AGGCTACCTTTATTTCTCTGACTGTGGTTGGCTGGAGACATGTCCCGCTTGGGCTG 401
 QY 121 ArgLysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGlnHisLeu--PheAlaG 140
 DB 402 CGCAAGATTTGGACAGAGCCATCTTGCATGATGTTTAGAACAA-AG-TCTTTGGCTC 459
 QY 140 LnhHisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLysGluGlnG 160
 DB 460 AGCACCGGAGTACCTGTGTGAGTGGCGCCCTCCCGTGGCTTGCATGAGAGAGCAGG 519
 QY 160 LysGluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGlnAlaGln 180
 DB 520 GAGAGTTCCTAGGCTGGCCGCTGGCTGGAGTGGCCAGATGGCTGTGAGCAGGCCAGC 579
 QY 180 TgProGlyGluLeuLeuLysThrVal 188
 DB 580 GCCCAGGAGACTGCTGAAGACGGTC 605

RESULT 6
 AA510093 512 bp mRNA linear EST 08-JUL-1997
 LOCUS AA510093
 DEFINITION v35c07.r1 Soares_mammary_gland_NbMMG Mus musculus cDNA clone
 IMAGE:863340 5' similar to gb:U32955 Mouse protein tyrosine kinase
 (MOUSE);, mRNA sequence.

ACCESSION AA510093
 VERSION AA510093.1 GI:2247947
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 512)
 REFERENCE
 AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
 Geisels,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,B.,
 Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
 Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and

TITLE Waterston,R.
 JOURNAL The Mashu-HHMI Mouse EST Project
 UNPUBLISHED (1996)
 COMMENT Contact: Marra M/Mouse EST project
 WashU-HHMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LLNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:507428
 Seq primer: -28m13 rev2 ET from Amersham
 High quality sequence stop: 373.

FEATURES
 source
 Location/Qualifiers
 1..512
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="IMAGE:863340"
 /clone_lib="Soares_mammary_gland_NbMMG"
 /sex="male"
 /russue_type="mammary gland"
 /dev_stage="4 weeks"
 /lab_host="DH10B"
 /note="Organ: mammary gland; Vector: pT7T3D-Pac (Pharmacia
) with a modified polylinker; Site:1: Not I; Site:2: Eco
 RI; 1st strand cDNA was primed with a Not I - oligo(dT)
 primer '15'
 TGTTCACATCTGTAAGTGGAGCGCGCCGACATGCTTTTCTTTTCTTTTCTTTT
 T 3'1; double-stranded cDNA was ligated to Eco RI
 adaptors (Pharmacia), digested with Not I and cloned into
 the Not I and Eco RI sites of the modified pT7T3 vector.
 RNA provided by Dr. Minoru Ko, Wayne State Univ. Library
 constructed and normalized by Bento Soares and M. Fatima
 Bonaldo."

BASE COUNT 90 a 171 c 156 g 94 t 1 others

ORIGIN

Alignment Scores:

Pred. No.:	Length:	Matches:
7.3e-97	512	110
Score: 110.00		
Percent Similarity: 100.00%	Conservative: 0	
Best local Similarity: 100.00%	Mismatches: 0	
Query Match: 10.01%	Indels: 0	
DB: 9	Gaps: 0	

US-09-397-967-16 (1-1099) x AA510093 (1-512)

QY 966 AlaLysLeuLeuProLeuGlyLysAspTyrTyrValValArgGluProGlyGlnSerPro 985
 DB 2 GCTAAGCTGTGCTGCCCTGGGAAAGAGACTACGTCGTGGCGGAGCTGGCCAAAGCCCC 61
 QY 986 IlePheTrpTyrAlaProGluSerLeuSerAspAsnIlePheSerArgGlnSerAspVal 1005
 DB 62 ATCTTTGTATGCCCCCGAGAGCCCTATCTGACACATCTTCTCCCGCAATCTAGCTG 121
 QY 1006 TrpSerPheGlyValValLeuTyrGluLeuPheThrTyrCysAspLysSerCysSerPro 1025
 DB 122 TGGACCTTCGAGTGGTGTGTTCAGCAGCTCTTCACTTCTGCGACAAAGACTGCGACCCA 181
 QY 1026 SerAlaGluPheLeuArgMetMetGlyProGluArgGluGlyProProLeuGlnSerArgLeu 1045
 DB 182 TCCGCTGATGTCCTGCGCATGATGGGCTGTAGGTGAAGGACCCCGCTGTGGCGCTC 241
 QY 1046 LeuGluLeuLeuAlaGluGlyLysArgLeuProProProProThrGlnVal 1065
 DB 242 CTGGAGCTGTGCGCAGAGAGCGGAGCGCTCCACACCACTCCACCTGCCACCGAGGTT 301
 QY 1066 GlnGluLeuMetGlnLeuCysThrAlaPro 1075
 DB 302 CAGGAGCTCATGACAGCTGTGCTGGGCGGCC 331

RESULT 7
BE916148 545 bp mRNA linear EST 29-SEP-2000
LOCUS 601666763p1.NCI.CGAP_Mam1 Mus musculus cDNA clone IMAGE:396903 5',
DEFINITION mRNA sequence.
ACCESSION BE916148
KEYWORDS GI:10416498
SOURCE EST.
ORGANISM house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 545)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLM9140 row: m column: 16
High quality sequence stop: 545.
Location/Qualifiers
1..545
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:396903"
/clone_lib="NCI.CGAP_Mam1"
/tissue_type="tumor, biopsy sample"
/dev_stage="10 months, virgin"
/lab_host="DH10B"
/note="Organ: mammary; Vector: PCMV-SPORT6; Site:1; Salt;
Site:2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"
BASE COUNT 93 a 176 c 166 g 110 t
ORIGIN
Alignment Scores:
Pred. No.: 3,29e-79 Length: 545
Score: 92.00 Matches: 92
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 8.37% Indels: 0
Gaps: 0
DB: 12
US-09-397-967-16 (1-1099) x BE916148 (1-545)
QY 374 GlyProIleThrLeuAspPheAlaIleHisLysLeuLysAlaAlaGlySerLeuProGly 393
DB 206 GCACCATCAGCTTAGACTTTCATCCACAGCTGAAGCGCGCTGCCCTCCAGGC 265
QY 394 ThrTrpIleLeuArgSerProGlnAspTrpAspSerPheLeuLeuThrAlaCysVal 413
DB 266 ACCATATATCTCCGCCGACGCCGAGCACTATGACAGCTTCTTACCCCTGCGTC 325
QY 414 GlnThrProLeuGlyProAspTrpLysGlyLysLeuIleArgGlnAspProSerGlyAla 433
DB 326 CAGACTCTCTTGGCCCGACACTACAAAGGGCTGCTCATCCGCCAGAGACCCGCGGGCT 385
QY 434 PheSerLeuValGlyLeuSerGlnProHisArgSerLeuArgGluLeuLeuAlaIaCys 453
DB 386 TTCTCCCTGGTGGCTCAGCCAGCCACACAAAGCTGCGGAGACTGCTTCACAGCTTC 445
QY 454 TrpAsnSerGlyLeuArgValAspGlyAlaAlaLeu 465
|||||

DB 446 TCGAATTCGTGGCTCGAGCTAGACGCTGCTGCCCTG 481
RESULT 8
B1557690 826 bp mRNA linear EST 05-SEP-2001
LOCUS 603237050f1.NCI.CGAP_Mam3 Mus musculus cDNA clone IMAGE:5289949 5',
DEFINITION mRNA sequence.
ACCESSION B1557690
KEYWORDS GI:15445004
SOURCE EST.
ORGANISM house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 826)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLM11733 row: d column: 14
High quality sequence stop: 760.
Location/Qualifiers
1..826
/organism="Mus musculus"
/strain="129,C57BL/6J,FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:5289949"
/clone_lib="NCI.CGAP_Mam3"
/tissue_type="tumor, gross tissue"
/dev_stage="10 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: PCMV-SPORT6; Site:1; Salt;
Site:2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigators
providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH
Reference for transgenic model: Xu et al., Nature Genetics
22, 37-43 (1999)."
BASE COUNT 147 a 252 c 240 g 187 t
ORIGIN
Alignment Scores:
Pred. No.: 4,45e-75 Length: 826
Score: 88.00 Matches: 208
Percent Similarity: 98.58% Conservative: 0
Best Local Similarity: 98.58% Mismatches: 3
Query Match: 8.01% Indels: 3
Gaps: 0
DB: 13
US-09-397-967-16 (1-1099) x B1557690 (1-826)
QY 1 MetaIaProProSerGluGluThrProLeuIleProGlnArgSerCysSerLeuSerSer 20
DB 37 ATGGACCTCCCAAGAGAGACACCTTGATCCCTGACGCGCTTGACGCTTCATCC 96
QY 21 SerGluAlaGlyAlaLeuHisValLeuLeuProProArgGlyProGlyProProGlnArg 40
DB 97 TCAGAGCGCAGAGCCCTGCTGCTCTCTCCCGGGACCTGGGCTCCGCCAGCCA 156
QY 41 LeuSerPheSerPheGlyAspTrpLeuAlaGluAspLeuCysValArgAlaAlaCysAla 60
DB 157 TTGTCTATTCTC-TTTGGGAGACTACTTGGCTGAGATTATGTGTGGAGACTCCCAAGGC 215
QY 61 CysGlyIleLeuProValTrpHisSerLeuPheAlaLeuAlaThrGluAspPheSerCys 80
DB 216 TGTGCATTCCTCCCTGTTATCATTCGCTTTCGCTGCGACATGAGAGACTTCTCTTCG 275
|||||

QY 81 TrpHeProSerHisIlePheCysIleGluAspValaSerThrGlnValLeuValTyr 100
 |||||||
 Db 276 TGGTTTCCCAACGACCATTTCTGCATAGAGACGTGGACACTCAAGCTTGGTCTAC 335
 |||||||
 QY 101 ArgLeuArgPheTyrPheProAspTrpPheGlyLeuGluThrCysHisArgPheGlyLeu 120
 |||||||
 Db 336 AGGTACGCTTTATTTCCCTGCAGCTGGTTTGCGCTGGAGACATGTCAACCGCTT-GGGGCTT 394
 |||||||
 QY 121 ArgTysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLeuPheAlaGln 140
 |||||||
 Db 395 CGCAAGA-ATTGACCAAGTGCATCTTGACTTACATGTTTAGAACATCTTTGGCTAG 453
 |||||||
 QY 141 HisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetTylGluGlnGly 160
 |||||||
 Db 454 CACCGCAGTGACCTGGTGTAGTGGGCGCTCCCGTGGGCTTACATGAAAGAGAGAGGA 513
 |||||||
 QY 161 GluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGlnAlaGlnArg 180
 |||||||
 Db 514 GAGTTCCTGACCTGGCGCTGCTGACTGGCCAGATGGCTGTGACAGAGCCAGCGC 573
 |||||||
 QY 181 ProGlyLeuLeuLeuLysThrValSerTyrLysAlaCysLeuProProSerLeuArgAsp 200
 |||||||
 Db 574 CAGAGAACGCTGCTGAAGACGCTGATTCACAAAGCTCTTCTGCGCCAGCGCTGGCAT 633
 |||||||
 QY 201 ValIleGlnGlnAsnPheValThrArgArg 211
 |||||||
 Db 634 GTGATCCAGGCGCAGAACTTGTGACACGACGAG 666
 |||||||
 RESULT 9
 LOCUS BF150250 675 bp mRNA linear EST 29-DEC-2000
 DEFINITION uy83609.y1 NCI_CGAP_Mam5 Mus musculus cDNA clone IMAGE:366184 5'
 similar to TR:P97423 P97423 JANTUS KINASE 3 ;, mRNA sequence.
 ACCESSION BF150250
 VERSION BF150250.1 GI:11031645
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 675)
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 JOURNAL Other ESTs: uy83609.x1
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapds-remail.nih.gov
 Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNLN at:
 image.lnl.gov/image/html/iresources.shtml
 MGI:1426952
 Seq primer: -40RP from Gibco
 High quality sequence stop: 390:
 FEATURES
 source Location/Qualifiers
 1. 675
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone IMAGE:366184
 /clone_lib="NCI_CGAP_Mam5"
 /tissue_type="tumor, gross tissue"
 /dev_stage="7 months"
 /lab_host="DH10B"
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: Salt;
 Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
 Library constructed by Life Technologies. Investigators

BASE COUNT 138 a 215 c 197 g 123 t 2 others
 ORIGIN
 Alignment Scores:
 Pred. No.: 2.62e-69 Length: 675
 Score: 82.00 Matches: 170
 Percent Similarity: 97.70% Conservative: 0
 Best local Similarity: 97.70% Mismatches: 3
 Query Match: 7.46% Indels: 4
 DB: 12 Gaps: 0
 US-09-397-967-16 (1-1099) x BF150250 (1-675)
 QY 717 GTPProAlaHisIleThrSerLeuGluProAlaLysTylSerPheTyrGluAspGln 736
 |||||||
 Db 37 GGACCGCCCAATCATTCTGCTGGAGCCCGCAAAAGCTGAAGTTTATGAGGAGCAG 96
 |||||||
 QY 737 GlyIleLeuProAlaLeuLysTrpThrGluLeuAlaGlyLeuIleThrGlnCysMetAla 756
 |||||||
 Db 97 GGACAGCTGCCCGCTCAATATGACAGACTGGGGGACTTATACACAGTGCATGGCG 156
 |||||||
 QY 757 TyrAspProGlyArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIle 776
 |||||||
 Db 157 TATGATCCTGGCGCGCGCCCTCTCGACTATCTCAAGACTCAAGCGCTCATTT 216
 |||||||
 QY 777 ThrSerAspTyrGluLeuLeuSerAspProThrProGlyIleProSerProArgAspGlu 796
 |||||||
 Db 217 ACATCAGATTACGAGCTCTCTCAAGACCCACACTGGCATCCGAGTCCCTCGAGATGAG 276
 |||||||
 QY 797 LeuCysValAlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGlnGlnArg 816
 |||||||
 Db 277 CTGTGCG--GT-GGCGCCCACTCTATGCTGCGCAGAGCCCGCATATTCGAGAGAGAGA 333
 |||||||
 QY 817 HisLeuLysTyrIleSerLeuLeuGlySerGlyAsnPheGlySerValGluLeuCysArg 836
 |||||||
 Db 334 CACCTTAAGTACATCTCTTCTGGGCAAGGCAACTTGGCAGCGTGGAGCTGTGCCG 393
 |||||||
 QY 837 TyrAspProLeuGlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGlnHisSer 856
 |||||||
 Db 394 TATGACCCCTGGGGGAGCAATATACGGGACCCTGTGGGAGTGAAGAACTACAGACAGC 453
 |||||||
 QY 857 Val-ProAspGlnGlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuHisSe 876
 |||||||
 Db 454 GG-GCCAGACACGACAGAGAGACTTCCACGCGGAGATTCAGATCTTAAGGCTCTGCACAG 512
 |||||||
 QY 876 RasPheIleValLysTyrArgGlyValSerTyrGlyPro 889
 |||||||
 Db 513 CGACTTCATCGTCAATACCGGGGAGTCACTATGGGCCA 552
 |||||||
 RESULT 10
 LOCUS BG870513 907 bp mRNA linear EST 29-MAY-2001
 DEFINITION 602791432F1 NCI_CGAP_S62 Mus musculus cDNA clone IMAGE:4922742 5',
 mRNA sequence.
 ACCESSION BG870513
 VERSION BG870513.1 GI:14221053
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 907)
 NIH-MGC http://mgc.ncbi.nlm.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapds-remail.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
 DNA Sequencing by: Incyte Genomics, Inc.

SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 478)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.
TITLE The WashU-HIMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Maria M/Mouse EST Project
WashU-HIMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MG1:274002
Seq primer: -28M13 rev2 from Amersham
High quality sequence stop: 463.
FEATURES Location/Qualifiers
1..478
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:457114"
/clone_1lb="Soares mouse placenta 4NDMP13.5 14.5"
/sex="unknown"
/tissue_type="placenta"
/dev_stage="adult"
/lab_host="DH10B"
/notes="Organ: placenta; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer; 15TGGTACCAATCTGAAGTCGGAGCGCCCGGAAATTTTTTTTTTTTTTTTTTTTTT
T 3]; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT 83 a 160 c 142 g 93 t
ORIGIN
Alignment Scores:
Pred. No.: 1.18e-61 Length: 478
Score: 74.00 Matches: 118
Best Similarity: 98.33% Conservative: 0
Percent Local Similarity: 98.33% Mismatches: 1
Query Match: 6.73% Indels: 2
Dbs: 9 Gaps: 0
US-09-397-967-16 (1-1099) x AA023670 (1-478)
Oy HisValIySIIeAlaAspPheGlyLeuAlaIySLeuLeuProLeuGlyIyAspTyrr 976
Db 3 CATGTGAAGATCGCGGACTTGGCTTCCCTAAGCTGTGCGCCCTGGGAAGACTACTAC 62
Oy 977 ValValAArgIuProGlyGlnSerProIlePheTyrIyAlaProGlnSerLeuSerAsp 996
Db 63 GTGGTCCCGAGACCTGGCCAAAGCCCATCTTTGTGTATGCCCGGAGTCCCTACTATAC 122
Oy 997 AsnIlePheSerAArgGlnSerAspValIrrPsePheGlyValValIleuTyrgIuLeuPhe 1016
Db 123 AACATCTTCCCGGCAATCTGACCTGTGGAGCTTCGGAGAGTGCTGTAGAGACTCTTC 182
Oy 1017 ThrTyrcysAspIySerCysSerProSerAlaGluPheLeuArg-MetMetGlyProG1 1036
Db 183 ACCTACTGCGACAGAGCTGAGAGCCCATCTCGCTGAGTTCCTGAG-CATGATGGGGCGCTGA 241

```

QY      1036  uaguglugglProProleucCysArgleuengluLeuLeuAgluglYArArleuPr 1056
          |||||
Db      242  GCGTGAAGAGACCCCCCTCTGCCCTCTCGAGCTGCTGGCAGAGGCGGCGCTCC 301
          |||||

QY      1056  opProProPrOthChySPProthrcInuValgIngluLeuMetgInLeuCYsTrPALAPro 1075
          |||||
Db      302  ACCACCTCCACCTGCGCCACCGAGGTTCAGAGGCTCAGCAGCGTGTGTGGGCCCC 359
          |||||

RESULT 13
LOCUS      AA881653
DEFINITION vx21c02.r1 Soares_thymus_2NDMT Mus musculus cDNA clone
            IMAGE:1265090 5' similar to gp:140172 Mus musculus JAK3 gene,
            complete cds (MOUSE);, mRNA sequence.
ACCESSION  AA881653
VERSION     AA881653.1 GI:2990963
KEYWORDS   EST.
SOURCE      house mouse.
            Mus musculus
ORGANISM   Mus musculus
REFERENCE  1 (bases 1 to 593)
AUTHORS    Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,D., Morris,M.,
            Scheinberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theisinger,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
TITLE       The WashU-HMI Mouse EST Project
JOURNAL    Unpublished (1996)
COMMENT    Contact: Maria M/Mouse EST Project
            WashU-HMI Mouse EST Project
            Washington University School of Medicine#
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:667642
            Seq primer: -28m13 rev2 ER from Amersham
            High quality sequence stop: 556.
            Location/Qualifiers
                1..593
                /organism="Mus musculus"
                /strain="C57BL/6J"
                /db_xref="taxon:10090"
                /clone="IMAGE:1265090"
                /clone_lib="Soares_thymus_2NDMT"
                /sex="male"
                /tissue_type="Thymus"
                /dev_stage="4 weeks"
                /lab_host="DH10B"
                /note="vector: pT73D-Pac (Pharmacia) with a modified
                polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                was primed with a Not I - o190(dt) primer [5',
                TGTACCAATCTGAGAGCGAGCGCGCGCTTTTATTTTTTTTTTTT
                3']; double-stranded cDNA was ligated to Eco RI adaptors
                (Pharmacia), digested with Not I and cloned into the Not I
                and Eco RI sites of the modified pT713 vector. RNA
                provided by Dr. Bertrand Jordan. Library went through two
                rounds of normalization, and was constructed by Bento
                Soares and M.Fatima Bonaldo."
BASE COUNT 106 a 197 c 166 g 124 t
ORIGIN
Alignment Scores:
Prid. No.: 1.45e-60 Length: 593
Score: 73.00 Matches: 117
Percent Similarity: 98.32% Conservative: 0
Best Local Similarity: 98.32% Mismatches: 1
Query Match: 6.64% Gaps: 2
DB: 9 Indels: 0

```

US-09-397-967-16 (1-1099) x AA881653 (1-593)

QY 958 VALIIVSIIAIAAAPHGGLYLEUAlALysLeuLeuProLeuGlyLysAspTYrTYrVal 977
 DB 15 GTGAAGATCCGCCACTTCGGCTCGCTAGCTGTGCCCCCTGGAGAAAGACTACTACCTG 74

QY 978 VALATGluProGluGlnSerProIlePheTYrTYrAlaProGluSerLeuSerAspAsn 997
 DB 75 GTCCCGAGACCTGGCCAAAGCCCATCTTTGGTATGGCCCGAGAGCTTATCTGCAAC 134

QY 998 IIEPHESerArgLInSerAspValTrpSerPheGlyValValLeuTYrGluLeuPheThr 1017
 DB 135 ATCTTCTCCCGCCAACTGACGTGAGCTTCGGAGTGGTGTGTGACAGCTCTTCACC 194

QY 1018 TYrCysAspLysSerCysSerProSerAlaGluPheLeuArg-MetMetGlyProGluArg 1037
 DB 195 TACTCGACAGACGCTGCGCCCATTCGCTGAGTTCTTGAG-CATGATGGGGCTGAGCG 253

QY 1037 GGLUGlyProProLeuGlyArgLeuLeuGluLeuLeuAlaGluGlyArgArgLeuProBr 1057
 DB 254 TGAAGAGACCCCGCTCTGCGCTCTGAGCTGTGAGGAGGCGGAGCGCTCCACC 313

QY 1057 OProProThrCysProThGluValGlnLeuLeuMetGlnLeuCysTrpAlaPro 1075
 DB 314 ACCTCCACCTGCGCCACCGAGGTTGAGAGCTCATGACGCTGTGCTGGCGCC 368

RESULT 14
 LOCUS BM935182 698 bp mRNA linear EST 13-MAR-2002
 DEFINITION UI-M-BH3-agz-f-08-0-UI.f1 NIH_BMAP_M.S4 Mus musculus cDNA clone
 ACCESSION BM935182
 VERSION UI-M-BH3-agz-f-08-0-UI 5', mRNA sequence.
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 698)
 Normalization and subtraction: two approaches to facilitate gene
 discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 COMMENT Contact: Chin, H
 National Institute of Mental Health
 6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
 20892-9643, USA
 Tel: 301 443 1706
 Fax: 301 443 9890
 Email: MES@emil.nih.gov
 CDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research
 Genetics (www.resgen.com).
 Seq primer: M13 REVERSE.
 FEATURES
 source location/Qualifiers
 1..698
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UI-M-BH3-agz-f-08-0-UI"
 /clone_lib="NIH_BMAP_M.S4"
 /dev_stage="27-32 days"
 /lab_host="DH10B (Life Technologies)"
 /note="Vector: pTZ19D-Pac (Pharmacia) with a modified
 polylinker; Site_1: Not I; Site_2: Eco RI; The
 NIH_BMAP_M.S4 library is a subtracted library of a series,
 ultimately derived from a mixture of individually tagged
 normalized libraries from ten regions of the mouse brain
 (cerebellum, brain stems, olfactory bulbs, hypothalamus,

cortex, amygdala, basal ganglia, pineal gland, striatum, hippocampus) after a series of subtractions to reduce the representation of cDNAs from which ESTs had already been generated. The following serially subtracted libraries were generated in this process: NIH_BMAP_M.S4, NIH_BMAP_M.S3.3, NIH_BMAP_M.S3.2, NIH_BMAP_M.S3.1, NIH_BMAP_M.S2, NIH_BMAP_M.S1. The subtracted library (NIH_BMAP_M.S4) was constructed as follows: PCR amplified cDNA inserts from NIH_BMAP_M.S3.3, NIH_BMAP_M.S3.2, and NIH_BMAP_M.S3.1 clones from which 3' ESTs had been derived was used as a driver in a hybridization with a pool of the NIH_BMAP_M.S3.3, NIH_BMAP_M.S3.2, and NIH_BMAP_M.S3.1 libraries in the form of single-stranded circles. The remaining single-stranded circles (subtracted library) was purified by hydroxyapatite column chromatography, converted to double-stranded circles and electroporated into DH10B bacteria (Life Technologies) to generate the NIH_BMAP_M.S4 library. This procedure has been previously described (Bonaldo, Lennon and Soares, Genome Research 6:791-806, 1996)"

BASE COUNT 132 a 211 c 215 g 138 t 2 others
 ORIGIN

Alignment Scores:
 Pred. No.: 1,66-58 Length: 698
 Score: 71.00 Matches: 183
 Percent Similarity: 95.31% Conservative: 0
 Best Local Similarity: 95.31% Mismatches: 3
 Query Match: 6.46% Indels: 9
 DB: 14 Gaps: 0

US-09-397-967-16 (1-1099) x BM935182 (1-698)

QY 801 GLYAlaGlnLeuTYrAlaCysGlnAspProAlaIlePheGlnGluArgHisLeuTYr 820
 DB 27 GGGGCCACCTGATGCTCTGCGAGGACCCCGCATGTTCAGAGACACCTTAAGTAC 86

QY 821 IIESerLeuGlnGlyLysLysAsnPheGlySerValGlnLeuCysArgTYrAspProLeu 840
 DB 87 ATCTCTTCTGCGCAAGGCAACTTTGGACGCTGAGCTGTGCGCTATGACCCCTG 146

QY 841 GLYAspAsnThrGlyProLeuValAlaValLysGlnLeuGlnHisSerVal-ProAsp 860
 DB 147 GGGGACATATACGGGACCCCTGCGAGGAGTGAACACTACAGACACGGG-GCCACACA 205

QY 860 nGlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuHisSerAspPheLeu 880
 DB 206 GCAGAGGACCTTCCACGCGGAGATTCAGATCTTAAGGCTCTGCACAGGACTTCATGCT 265

QY 880 ILYSTYrArgGlyValSerTYrGlyProGluArgGlnSerLeuArgValMetGluTYr 900
 DB 266 CAAGTACCGGGAGTACGATATGCGGACAGGCTCGGCTGTGGTATGGAGTA 325

QY 900 ILEuProSerGlyCysLeuArgAspLeu-LeuGlnArgHis-----ArgGlyLeuHisTYr 918
 DB 326 CTGCGCCACGCGGCTGCTGGAGAA-CTTCTGACAGGCGCATCCCGGCGCG-CGACACAC 383

QY 919 ASPArgLeuLeuLeuPheAlaTrpGlnIleCysLysGlyMetGluTYrLeuGlyAlaArg 938
 DB 384 GACCGCTACTGCTGTTCGCTGGAGATCTGCAAGGCGCATGAGTAACCTGGGCGCGCC 443

QY 939 ARGCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGlnSerGluAlaHisVal 958
 DB 444 CGCTGGGTACACGTCACGTCGCGCGGCAACATCTTGTTGAGAGCGGAGCTCATGTG 503

QY 959 LYSIIAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlnLysAspTYrTYrValAla 978
 DB 504 AAGATCGCGGACTTCGGCTCGCTAAGCTCTGCCCTGGAGAAAGACTACTACGTGTC 563

QY 979 ARGGluProGluGlnSerProIle 986
 DB 564 CGGAGCTGCGCAAGGCCCATC 587


```

RESULT 15
BF181053
LOCUS
DEFINITION BF181053 344 bp mRNA linear EST 31-OCT-2000
601805105f1 NCL.CGAP_Mam5 Mus musculus cDNA clone IMAGE:4035936 5',
mRNA sequence.
ACCESSION BF181053
VERSION BF181053.1 GI:11059195
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE NIH-MGC
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-rt@mail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM9310 row: j column: 01
High quality sequence stop: 344.
Location/Qualifiers
1..344
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone_image="IMAGE:4035936"
/clone_lib="NCI CGAP Mam5"
/tissue_type="tumor_gross tissue"
/dev_stage="7 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigators
providing samples: Lothar Hennighausen/Robin Humphreys,
NIH"
BASE COUNT 73 a 95 c 108 g 68 t
ORIGIN
Alignment Scores:
Pred. No.: 6,27e-57 Length: 344
Score: 69.00 Matches: 113
Percent Similarity: 97.41% Conservative: 0
Best Local Similarity: 97.41% Mismatches: 2
Query Match: 6.28% Indels: 3
DB: 12 Gaps: 0
US-09-397-967-16 (1-1099) x BF181053 (1-344)
QY 646 ArgLysValLeuLeuAlaArgGlyGlyAspGlyAsnProporphylLeuSer 665
Db 1 CGGAAGGCTGCTCGCTCGTGAAGGGGAGTGAATGCACCTTCATTAACTAGT 60
QY 666 AspProGlyValSerProThrValLeuSerLeuLettLeuThraPargIleProTyr 685
Db 61 GATCTGTGTGTCAGTCCACAGCTGTCTAGCCTGGAATGCTCACGACAGATVACCTGC 120
QY 686 ValAlaProGlyLeuGlnGluAlaGlnThrLeuGlyLeuGlnAlaAspLysTyrGly 705
Db 121 GTGGCCCCCAAGATGTCACAGAGAGCTCAGACACTGCTTGGAGGCTGACAAAGTGGGGC 180
QY 706 PheGlyAlaThrThrTyrGluValPheGlnArg-GlyProAlaHisIleThrSerLeuGln 725
Db 181 TTGGAGCCCAACAGTGGAGGAGTGT-T-CACGGGGGAGCCGCCACATCACCTCGCTGGA 239
QY 725 -uProAlaLysLysLeuLysPheTyrGluAspGlnGlyLeuProAlaLeuLysTyrPth 745
|||||

```

```

Db 240 GCCCGCAAAAAGCTGAAGTTCTATGAGGACA-GGACAGCTGCCCTCTCAATGAGC 298
QY 745 gLLeuAlaGlyLeuIleThrGlnCysMetaLarTyraProGly 760
Db 299 AGAAGTGGCGGACTTATCACACAGTGCATGCGCTATGATCTCTGCC 344
RESULT 16
AM141115
LOCUS
DEFINITION AM141115 436 bp mRNA linear EST 30-OCT-1999
EST291147 Normalized rat embryo, Bento Soares Rattus sp. cDNA clone
RG1BC74 5' end similar to protein-tyrosine kinase jak3, mRNA
sequence.
ACCESSION AM141115
VERSION AM141115.1 GI:6160891
KEYWORDS EST.
SOURCE Rattus sp.
ORGANISM Rattus sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
AUTHORS 1 (bases 1 to 436)
Lee,N.H., Glodex,A., Chandra,I., Mason,T.M., Quackenbush,J.,
Kerlavage,A.R. and Adams,M.D.
TITLE Rat Genome Project: Generation of a Rat EST (RESt) Catalog & Rat
Gene Index
JOURNAL Unpublished (1998)
COMMENT Contact: Lee, NH
The Institute for Genomic Research
9712, Medical Center Drive, Rockville, MD 20850, USA
Tel: (301)-838-3529
Fax: (301)-838-0208
Email: nhlee@igr.org
For clone availability, additional sequence and expression
information related to this EST please check the TIGR Rat Gene
Index (http://www.igr.org/tid/ygi/rat.html). To order a clone
contact the ATCC (http://www.atcc.org/atcc.html).
Seq primer: M13 Reverse.
Location/Qualifiers
1..436
/organism="Rattus sp."
/db_xref="taxon:10118"
/clone="RG1BC74"
/clone_lib="Normalized rat embryo, Bento Soares"
/dev_stage="embryo 8, 12, 18 dpc"
/note="Vector: pT7mpac; Site_1: EcoRI; Site_2: NotI"
BASE COUNT 72 a 134 c 108 g 122 t
ORIGIN
Alignment Scores:
Pred. No.: 6.84e-53 Length: 436
Score: 65.00 Matches: 65
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.91% Indels: 0
DB: 10 Gaps: 0
US-09-397-967-16 (1-1099) x AM141115 (1-436)
QY 13 GlnArgSerCysSerLeuSerSerGluAlaGlyAlaLeuHisValLeuLeuProPro 32
Db 61 CAGGCTCTTGTATACCTCTATCTCAGAAAGCAGGACCTCGATGCTGCTTCCCTCC 120
QY 33 ArgGlyProGlyProProGlnArgLeuSerPheSerPheGlyAspTyrLeuAlaGluAsp 52
Db 121 CGGGAGACTGGGCTCCCGACGACACTGCTATCTCTTTGGGAGCTACTGGCTGAGAT 180
QY 53 LeuGlyValArgAlaAlaLysAlaCysGlyIleLeuProValTyrHisSerLeuPheAla 72
Db 181 CTGTGTGTGGGGCTGCAAGAGCTGTGGCATCTGCGCCCTTATATCATTCGCTTCGCT 240
QY 73 LeuAlaThrGluAsp 77
|||||
Db 241 CTGGCCACGAGAGC 255

```


US-09-397-967-16 (1-1099) x BG800853 (1-559)

OY 427 ArgGlnAspProSerGlyAlaPheSerLeuValGlyLeuSerGlnProHisArgSerLeu 446
|||||
DB 37 CGGACAGAGCCAGCGGGGTTCTCCCTGGTTGGCTCAGTACGCCACAGAACCTTG 96
|||||
OY 447 ArgGlnLeuLeuAlaAlaLacYTrpAsnSerGlyLeuArgValAspGlyAlaAlaLeuTrp 466
|||||
DB 97 CGGAGACTCTGGACACGCTGGGAATCTGGGCTCGAGTACGAGTGGCTGCTCCGTA- 155
|||||
OY 467 -LeuThrSerCysAlaProArgProGlyGlySerAsnLeuIleValValArg 486
|||||
DB 156 CCTAACATCTCTGCTGGCTGCCAGACCCAGAGAAAGTCAATTGATGCTGGTGGAG 215
|||||
OY 486 gclYcysAsn-ProAlaProAlaProGlyCysSerProSerCysAlaLeuThrGln 506
|||||
DB 216 GGGCTGTGAC-CCCGCGCTGCGCCCTGGCTGCTCCCGTCTGCTGCTGCGTACACAC 274
|||||
OY 506 euserPheHisThrIleProThrAspSerLeuGluTrpHisGlnLeuGlyHisGly 526
|||||
DB 275 TGAGCTCCACACAAATCCACGACAGCCTGAGTGGCACAGAACCTGGGTCAAGGTT 334
|||||
OY 526 erPheThrIlePheArgGlySer-ArgArgGluValAlaAspGlyGluThrHisAsp 545
|||||
DB 335 CTTTACCAAGATCTTCCGTGGCCG-CAGGCGGAGGTGCTGATGCTAGACACATGAC 393
|||||
OY 546 SerGluValLeuLeuLysValMetAspSerArgHisArgAsnCysMetGluSerPheLeu 565
|||||
DB 394 TCGGAAGTCTCTCAGAGTCAATGAGCTCCAGACATCGGAACCTGATGAGCTTTCTG 453
|||||
OY 566 GluAlaAlaSerLeuMetSerGlnValSerTrpProHisLeuValLeuLeuHisGlyAl 585
|||||
DB 454 GAAAGCCGAGAGTTATGAGCCCAAGTATCTTACCGCACCTGGTGTACTGACAGCGCTG 513
|||||
OY 586 CysMetAlaGlyAspSerIleMetValGlnGlu 596
|||||
DB 514 TGCATGCTGAGACAGCATCATGTGCAGGAA 546
|||||
RESULT 19
LOCUS B1153609 358 bp mRNA linear EST 05-JUL-2001
DEFINITION 60287117BF1 NCI-CGAP_Mam2 Mus musculus cDNA clone IMAGE:5003274 5',
mRNA sequence.
B1153609
B1153609.1 GI:14613610
EST.
ACCESSION B1153609.1
VERSION B1153609
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 358)
TITLE NIH-MGC http://mgi.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-femail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
plate: LLM1039 row: k column: 19
High quality sequence stop: 358.
Location/Qualifiers
1. 358
/organism="Mus musculus"
/strain="FVB/N-3"
/db_xref="taxon:10090"
/clone="IMAGE:5003274"
/clone_lid="NCI-CGAP_Mam2"
/tissue_type="tumor, biopsy sample"

/dev stage="5 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-Sport6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"

BASE COUNT 77 a 100 c 110 g 71 t

ORIGIN

Alignment Scores:

Pred. No.:	4,25e-48	Length:	358
Score:	60.00	Matches:	117
Percent Similarity:	96.69%	Conservative:	0
Best Local Similarity:	96.69%	Mismatches:	2
Query Match:	5.46%	Indels:	4
DB:	13	Gaps:	0

US-09-397-967-16 (1-1099) x B1153609 (1-358)

OY 642 AsnValSerAlaArgLysValLeuLeuAlaArgGlnGlyAspGlyAsnProPro-Ph 661
|||||
DB 2 AACGCTCAGCAGGAGAGTGTCTGCTGCTGAGGGGGGATGGATACACTTTT 61
|||||
OY 661 ellleYsLeuSerAspProGlyValSerProThrValLeuSerLeuGluMetLeuThrAs 681
|||||
DB 62 CATTAAGCTGAGTATCTCTGTGTCAGTCCCACTGCTGAGCCTGGAATGCTCACCGA 121
|||||
OY 681 parGtleProTrpValAlaProGlyCysLeuGlnGluValGlnThrLeuCysLeuGluAl 701
|||||
DB 122 CAGATACCTGGTGGTGGCCCGCCCAATGCTCCAGAGGCTCAGACATCTGCTGGAGGC 181
|||||
OY 701 aAspLysTrpGlyPheGlyAlaThrThrTrpGluValPheGlnArg-GlyProAlaHisI 721
|||||
DB 182 TGACAGAGTGGGCTTTGGAGCACACAGTGGAGGTGTT-CAGCGGGGACCCGCCACA 240
|||||
OY 721 leThrSerLeuGlnProAlaLysLysLeuLysPheTyGluAspGlnGlnLeuProA 741
|||||
DB 241 TCACCTCGCTGAGCCGCCCAAAAGCTGAAGTTCTATGAGAGACA-GGACAGCTGCCG 299
|||||
OY 741 laleLysTrpThrGluLeuAlaGlyLeuIleThrGlnCysMetAlaTrpAspProGly 760
|||||
DB 300 CTCCTCAATGAGCAGAACTGGCGGAGCTTATCAGACACATGCATGCGCTATCATCTGTC 358
|||||
RESULT 20
LOCUS AW742526 621 bp mRNA linear EST 07-SEP-2000
DEFINITION up58e12.y1 Soares mouse NMIE Mus musculus cDNA clone IMAGE:2780302
5', similar to gb:U40172 Mus musculus JAK3 gene, complete cds (MOUSE
); mRNA sequence.
AW742526
AW742526.1 GI:7654317
EST.
ACCESSION AW742526.1
VERSION AW742526
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 621)
TITLE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
JOURNAL National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
COMMENT Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-femail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Seq primer: -40BP from Gibco
High quality sequence stop: 471.
Location/Qualifiers
1. 621
/organism="Mus musculus"
/strain="C3H x 101 F1"
/db_xref="taxon:10090"
/clone="IMAGE:2780302"


```

VERSION      AA755769.1  GI:2802967
KEYWORDS     EST.
SOURCE       house mouse.
ORGANISM     Mus musculus
REFERENCE    Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
AUTHORS      1 (bases 1 to 867)
             Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
             Geisels, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
             Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
             Treising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
             Waterston, R.
TITLE        The WashU-HMI Mouse EST Project
JOURNAL      Unpublished (1996)
COMMENT      Contact: Marra M/Mouse EST Project
             WashU-HMI Mouse EST Project
             Washington University School of Medicine
             4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
             Tel: 314 286 1800
             Fax: 314 286 1810
             Email: mouseest@watson.wustl.edu
             This clone is available royalty-free through LLNL; contact the
             IMAGE Consortium (info@image.llnl.gov) for further information.
             MGI:649972
             Seq primer: -28m13 rev1 ET from Amersham
             High quality sequence stop: 412.
             Location/Qualifiers
               1..867
               /organism="Mus musculus"
               /strain="NIH Swiss"
               /db_xref="taxon:10090"
               /clone_image="1224380"
               /clone_lib="Stratagene mouse heart (#937316)"
               /sex="pooled"
               /tissue_type="heart"
               /dev_stage="13 day embryos"
               /lab_host="SOLR (kanamycin resistant)"
               /note="Organ: heart; Vector: pBluescript SK-; Site:1:
               EcoRI; Site:2: XhoI; Cloned unidirectionally. Primer:
               Oligo dt. 93 pooled NIH/Swiss 13 day embryo hearts.
               Average insert size: 1.0 kb; Uni-ZAP XR Vector; -5'
               adaptor sequence: 5' GAATTCGGCAGCAG 3' -3' adaptor
               sequence: 5' CTCGAGCTTTTCTTTTCTTTT 3'"
BASE COUNT   179 a      239 c      292 g      157 t
ORIGIN
Alignment Scores:
Pred. No.:    7,87e-39      Length:      867
Score:         51.00        Matches:      78
Percent Similarity: 97.50%   Conservative: 0
Best Local Similarity: 97.50% Mismatches:    2
Query Match:    4.64%      Indels:        2
DB:              9          Gaps:          0
US-09-397-967-16 (1-1099) x AA755769 (1-867)
QY      991 ProGUlSerLeuSerAspAsnIlePheSerArgInSerAspValIrrPserPheGlyVal 1010
      |||
Db      454 CCGAGATCCCTATCTGACACATCTTCTCCGCCCATCTGACGTGTGGAGTGTGGAGTG 395
QY      1011 ValLeuTyrlGluLeuPheTrhTyrcysAspIlySerSsrProserAlaGluPheLeu 1030
      |||
Db      394 GTGTGTGACAGAGCTTCTCACCCTACTGCGACAGAGCTGCAGCCCATCGCTGAGTTCCTG 335
QY      1031 ArgMetMetGlyProGUATrgGluGlyProProLeuCySarGleuLeuGluLeuAla 1050
      |||
Db      334 CGCATGATGGGGCTGTGAGCGTGAAGACCCCGTC-TCGCCCTCTCTGAGACTCTGGCA 276
QY      1051 GluGlyArgArgLeuProProProThrcysProThrcysGluValGluGluMetGln 1070
      |||
Db      275 GA-GGGCGAGCGCTCCACACACCTCCACCTGCCCGACCGAGGTTTCAAGAGCTATCGAG 217
RESULT 23

```

```

BE753801
LOCUS       BE753801          543 bp      mRNA      linear      EST 25-APR-2001
DEFINITION 206963 MARC 2BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION  BE753801
VERSION    BE753801.1  GI:10167793
KEYWORDS   EST.
SOURCE     cow.
ORGANISM   Bos taurus
REFERENCE   Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS     Eukaryota; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
             Bovidae; Bovinae; Bos.
             1 (bases 1 to 543)
             Smith, T.P.L., Grosse, W.M., Freking, B.A., Roberts, A.J., Stone, R.T.,
             Casas, E., Wray, J.E., White, J., Cho, J., Fahrnenkrug, S.C., Bennett,
             G.L., Heaton, M.P., Laegreid, W.W., Rohrer, G.A., Chitko-McKown, C.G.,
             Pette, G., Holt, I., Karamycheva, S., Liang, F., Quackenbush, J. and
             Keefe, J.W.
TITLE       Sequence evaluation of four pooled-tissue normalized bovine cDNA
JOURNAL     Libraries and construction of a gene index for cattle
MEDLINE    Genome Res. 11 (4), 626-630 (2001)
COMMENT     Contact: Smith TPL
             USDA, ARS, US Meat Animal Research Center
             PO Box 166, Clay Center, NE 68933-0166, USA
             Tel: 402 762 4366
             Fax: 402 762 4390
             Email: smith@email.marc.usda.gov
             Single pass sequencing. Bases called and alt trimmed with phred
             v0.980904.e. Vector identified by cross_match with the -mnscore 18
             and -mismatch 12 options.
             PCR PRIMERS
             FORWARD: AGGAACAGCTATGACCAT
             BACKWARD: GTTTCCTCAGTCACGACG
             Plate: 49 row: M column: 16
             Seq primer: ATTAGGTGACACTATAG.
             Location/Qualifiers
               1..543
               /organism="Bos taurus"
               /db_xref="taxon:9913"
               /clone_lib="MARC 2BOV"
               /tissue_type="pooled"
               /lab_host="DH10B"
               /note="Vector: pCMV SPORT6; Site:1: NotI; Site:2: SalI;
               Library made from pooled tissue from testis, thymus,
               semitendinosus muscle, longissimus muscle, pancreas,
               adrenal, and endometrium."
BASE COUNT   100 a      182 c      158 g      103 t
ORIGIN
Alignment Scores:
Pred. No.:    4.1e-37      Length:      543
Score:         49.00        Matches:      49
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches:    0
Query Match:    4.46%      Indels:        0
DB:              12          Gaps:          0
US-09-397-967-16 (1-1099) x BE753801 (1-543)
QY      860 GInGlnArgAspPheGlnArgGluIleGlnIleLeuAlaLeuHisSerAspPheIle 879
      |||
Db      12 CAGCAGAGGGACCTTCAGCGGGAGATCCAGATCTCTCAAGCCCTCCACAGACTTATT 71
QY      880 ValIlySTyrrArgGlyValSerTyrrGlyProGUlArgInSerLeuArgLeuValMetGlu 899
      |||
Db      72 GTCAAGTACCGGGGTGTACGCTATGCCCCAGGCCGCCAGAGCTTGGTATGAGAG 131
QY      900 TyrlLeuProSerGlyCySLeuArgAsp 908
      |||
Db      132 TATCTGCCCATGTGCTGCTGCGCGAC 158
RESULT 24
BI339676

```

LOCUS BI339676 557 bp mRNA linear EST 30-JUL-2001
 DEFINITION 364724 MARC 2Pig Sus scrofa cDNA 5', mRNA sequence.
 ACCESSION BI339676
 VERSION BI339676.1 GI:15032955
 KEYWORDS EST.
 SOURCE Pig.
 ORGANISM Sus scrofa
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
 REFERENCE 1 (bases 1 to 557)
 Fahrnkung,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E.,
 Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Laegreid,W.M.
 and Keeler,J.W.
 Design and use of two pooled tissue normalized cDNA libraries for
 EST discovery in swine
 Unpublished (2000)
 CONTACT: Smith TPL
 JOURNAL USDA, ARS, US Meat Animal Research Center
 PO Box 166, Clay Center, NE 68933-0166, USA
 Tel: 402 762 4366
 Fax: 402 762 4390
 Email: smith@email.marc.usda.gov
 Single pass sequencing. Bases called and alt_trimmed with phred
 v0.980904.e. Vector identified by cross_match with the -minscore 18
 and -mismatch 12 options.
 PCR PRIMERS
 FORWARD: AGGAACAGCTATGACCAT
 BACKWARD: GTTTCCTCAGTCACGACG
 Plate: 102 row: E column: 17
 Seq primer: ATTAGGTGACACTATAG.
 FEATURES
 source
 1..557
 /organism="Sus scrofa"
 /db_xref="taxon:9823"
 /clone_lib="MARC 2Pig"
 /tissue_type="pooled"
 /lab_host="DH10B"
 /note="Vector: pCMV SPORT6; Site_1: NotI; Site_2: SalI;
 library made from pooled tissue from testis, ovary,
 endometrium, hypothalamus, pituitary, and placenta."
 BASE COUNT 115 a 185 c 159 g 98 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 4,22e-37 Length: 557
 Score: 49.00 Matches: 49
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 4.46% Indels: 0
 DB: 13 Gaps: 0
 US-09-397-967-16 (1-1099) x BI339676 (1-557)
 QY 860 GlnGlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuHisSerAspPheLe 879
 DB 238 CACACAGAGGACTTCAGCGGGGAGATCCATTAAGCCCTCCACAGCAGCACTTCATT 297
 QY 880 ValLysTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuArgLeuValMetGlu 899
 DB 298 GTCAAGTACCGGGGTCTACGATATGCCCAAGGCCCAAGAGATCTGGGCTGTGATGAA 357
 QY 900 TyrLeuProSerGlyCysLeuArgAsp 908
 DB 358 TATCTGCCAAGCGGCTGCTCGCGGAC 384
 RESULT 25
 LOCUS BF705690 544 bp mRNA linear EST 29-DEC-2000
 DEFINITION 243606 MARC 2Pig Sus scrofa cDNA 5', mRNA sequence.
 ACCESSION BF705690
 VERSION BF705690.1 GI:11997351
 KEYWORDS EST.
 SOURCE Pig.

ORGANISM Sus scrofa
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
 REFERENCE 1 (bases 1 to 544)
 Fahrnkung,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E.,
 Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Laegreid,W.M.
 and Keeler,J.W.
 Design and use of two pooled tissue normalized cDNA libraries for
 EST discovery in swine
 Unpublished (2000)
 CONTACT: Smith TPL
 JOURNAL USDA, ARS, US Meat Animal Research Center
 PO Box 166, Clay Center, NE 68933-0166, USA
 Tel: 402 762 4366
 Fax: 402 762 4390
 Email: smith@email.marc.usda.gov
 Single pass sequencing. Bases called and alt_trimmed with phred
 v0.980904.e. Vector identified by cross_match with the -minscore 18
 and -mismatch 12 options.
 PCR PRIMERS
 FORWARD: AGGAACAGCTATGACCAT
 BACKWARD: GTTTCCTCAGTCACGACG
 Plate: 70 row: D column: 5
 Seq primer: ATTAGGTGACACTATAG.
 FEATURES
 source
 1..544
 /organism="Sus scrofa"
 /db_xref="taxon:9823"
 /clone_lib="MARC 2Pig"
 /tissue_type="pooled"
 /lab_host="DH10B"
 /note="Vector: pCMV SPORT6; Site_1: NotI; Site_2: SalI;
 library made from pooled tissue from testis, ovary,
 endometrium, hypothalamus, pituitary, and placenta."
 BASE COUNT 96 a 181 c 156 g 111 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 3,55e-34 Length: 544
 Score: 46.00 Matches: 46
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 4.19% Indels: 0
 DB: 12 Gaps: 0
 US-09-397-967-16 (1-1099) x BF705690 (1-544)
 QY 973 LysAspTyrTyrValValArgGlnProGlyGlnSerProIlePheTyrTyrAlaProGlu 992
 DB 175 AAAGACTACTATGTGTCTCCGCGACCCCGCCAGAGCCCATCTCTGTATGCCCGGAG 234
 QY 993 SerLeuSerAspAsnIlePheSerArgGlnSerAspValTyrPheGlyValValLeu 1012
 DB 235 TCCCTCTAGACATATCTTCTCCGCGAGTGGATGTCTGAGCTTGGGGGTCTGCTG 294
 QY 1013 TyrGlnLeuPheTyrTyr 1018
 DB 295 TATGAGCTCTTACACCTAC 312
 RESULT 26
 LOCUS BF193178 562 bp mRNA linear EST 02-NOV-2000
 DEFINITION 244670 MARC 2Pig Sus scrofa cDNA 5', mRNA sequence.
 ACCESSION BF193178
 VERSION BF193178.1 GI:11076547
 KEYWORDS EST.
 SOURCE Pig.
 ORGANISM Sus scrofa
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
 REFERENCE 1 (bases 1 to 562)
 Fahrnkung,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E.,
 Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Laegreid,W.M.

TITLE design and use of two pooled tissue normalized cDNA libraries for EST discovery in swine

JOURNAL Unpublished (2000)

COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt trimmed with phred v0.980904.e. Vector identified by cross_match with the -mnscore 18 and -minmatch 12 options.

FEATURES

source

1. 562
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone_lib="MARC_2Pig"
/tissue_type="pooled"
/lab_host="DH10B"
/note="Vector: PCMV SPORT6; Site_1: NotI; Site_2: SalI; library made from pooled tissue from testis, ovary, endometrium, hypothalamus, pituitary, and placenta."

BASE COUNT 97 a 188 c 167 g 110 t

ORIGIN

Alignment Scores:

Pred. No.:	Length:	Matches:
46.00	562	46
Score: 100.00%	Conservative: 0	
Percent Similarity: 100.00%	Mismatches: 0	
Best local Similarity: 4.19%	Indels: 0	
Query Match: 12	Gaps: 0	

DB:

US-09-397-967-16 (1-1099) x BF193178 (1-562)

QY 973 LYSAPTYRTYRVALAARGLUPROGLYINSERPROILEPHERTRYALAPROGLU 992
|||||

Db 186 AAGACTACTATGTGTGCGCGAGCGCGCCAGCCCATCTTGATATGCCCGGAG 245
|||||

QY 993 SERLEUSERASPANLLEPHESEARARGLINSERAPVALTRPSEPHEDLYVALLEU 1012
|||||

Db 246 TCCCTCTCAGACAATATCTTCGCGCGAGTGATGTGGAGCTTCGGGCTGCTG 305
|||||

QY 1013 TYRGLULEUPHERTRYR 1018
|||||

Db 306 TATGAGCTTCTCACCTAC 323
|||||

RESULT 27
A2956543 659 bp DNA linear GSS 27-APR-2001
LOCUS A2956543
DEFINITION 2M0223004F Mouse 10kb plasmid UUGC2M library Mus musculus genomic
clone UUGC2M0223004 F, DNA sequence.

ACCESSION A2956543
VERSION A2956543.1 GI:13827770
KEYWORDS GSS.
SOURCE house musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 659)
Dunn,D., Aoyagi,A., Barber,M., Beacom,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished (2000)

COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0223 row: 0 column: 04
Seq primer: GGTGTAAACGACGCGCCACT
Class: plasmid ends
High quality sequence stop: 659.

FEATURES

source

1. 659
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone_lib="UUGC2M0223004"
/clone_lib="Mouse 10kb plasmid UUGC2M library"
/sex="Female"
/lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PMD420V; Purified genomic DNA from M. musculus C57BL/6J (female) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The ligated DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PMD42 (g1473214.19b/AF129072.1), a copy number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 155 a 217 c 178 g 109 t

ORIGIN

Alignment Scores:

Pred. No.:	Length:	Matches:
45.00	659	45
Score: 100.00%	Conservative: 0	
Percent Similarity: 100.00%	Mismatches: 0	
Best local Similarity: 4.09%	Indels: 0	
Query Match: 17	Gaps: 0	

DB:

US-09-397-967-16 (1-1099) x A2956543 (1-659)

QY 325 GUAlAGLUPHeproGLYLeuProGLuAlaLeuSerPheValAlaLeuValAspGLYTYR 344
|||||

Db 452 GAAGCGGAGTTTCCGGGGCTGCTGAGCGCTGCTTCGCGCCCTGGATGGGTAC 393
|||||

QY 345 PHEARGLEULIECYASPSERARGHISTYRPHCYSYSGUVALALAPROPRARGLEU 364
|||||

Db 392 TTCGCCCTGATCTCGACTCCAGGCAATATTTCGCAAGAGGTGGCGCCCGGCTG 333
|||||

QY 365 LEUGLUGLUGLUALA 369
|||||

Db 332 CTGAGAGAGAGGCG 318
|||||

RESULT 28
B1540725 417 bp mRNA linear EST 30-AUG-2001
LOCUS B1540725
DEFINITION 454105 MARC 1BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION B1540725
VERSION B1540725.1 GI:15381837
KEYWORDS EST.

SOURCE	cow.	
ORGANISM	Bos taurus	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.	
AUTHORS	1 (bases 1 to 417) Smith,T.P.L., Grosse,M.W., Freking,B.A., Roberts,A.J., Stone,R.T., Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-Mckown,C.G., Pereira,G., Holt,L., Karameycheva,S., Liang,F., Quackenbush,J. and Keeler,J.W.	
TITLE	Sequence evaluation of four pooled-tissue normalized bovine cDNA libraries and construction of a gene index for cattle	
JOURNAL	Genome Res. 11 (4), 626-630 (2001)	
MEDLINE	21180013	
COMMENT	Contact: Smith TPL USDA, ARS, US Meat Animal Research Center PO Box 166, Clay Center, NE 68933-0166, USA Tel: 402 762 4366 Fax: 402 762 4390 Email: smith@email.marc.usda.gov Single pass sequencing. Bases called and alt_trimmed with phred v0.980904.e. Vector identified by cross_match with the -mismatch 12 options. PCR Primers FORWARD: AGGAACACGATGATGACCAT BACKWARD: GTTTCCTCCAGTCACGAG Plate: 118 row: P column: 4 Seq primer: ATTGAGGACGACATATG. Location/Qualifiers 1..417 /organism="Bos taurus" /db_xref="taxon:9913" /clone_lib="MARC 1BCV" /tissue_type="pooled" /lab_host="BDH10B" /note="vector: pCMV SPORT6, Site_1: NotI, Site_2: SalI; Library made from pooled tissue from lymph node, ovary, fat, hypothalamus, and pituitary." BASE COUNT 74 a 138 c 121 g 83 t 1 others ORIGIN	
FEATURES	Alignment Scores: Pred. No.: 2.35e-32 Length: 417 Score: 44.00 Matches: 44 Percent Similarity: 100.00% Conservative: 0 Best Local Similarity: 100.00% Mismatches: 0 Query Match: 4.00% Indels: 0 DB: 13 Gaps: 0	
US-09-397-967-16 (1-1099) x B1540725 (1-417)		
QY 975 TyrTyrValValAlaGluLysProGlyGlnSerProIlePheThrPtyrAlaProGluSerLeu 994 112 TACTACGCGTGGTCGCCGAGCCAGCCAGCCCATCTTCTGTATGCCCGCATCTCCCTC 171 		
QY 995 SerAspAniIlePheSerArgGlnSerAspAlaItrPsePheGlyValValLeuTyrGlu 1014 Db 172 TCGGACACACATCTTCTCGCCGACGATGTCTGTGAGCTTCGGGGTCTGTCTAGAG 231 		
QY 1015 LeuPheThrTyr 1018 		
Db 232 CTCCTTCACCTTAC 243		
RESULT 29		
AA020387	441 bp	linear
LOCUS	AA020387	mrna
DEFINITION	mb60c03.r1 Soares mouse placenta 4abmp13.5 14.5 Mus musculus cDNA clone IMAGE:455332.5' similar to gb:U40172 Mus musculus JAK3 gene, complete cds (MUSEF), mRNA sequence.	
ACCESSION	AA020387	
VERSION	AA020387.1	GI:1484169
FEATURES	EST.	

SOURCE	mouse musculus.
ORGANISM	Mus musculus.
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 441)
AUTHORS	Marral,M., Hillier,L., Allen,M., Bowles,M., Dierich,N., Dubouque,T., Geisler,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,S., Soares,B., Wilson,R. and Waterston,R.
TITLE	The WashU-HMI Mouse EST Project
JOURNAL	Unpublished (1996)
COMMENT	Contact: Marra M/Mouse EST Project WashU-HMI Mouse EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: mouseest@watson.wustl.edu This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information. MGI:272220
FEATURES	Putative full length read vector to vector length is 448 Seq primer: -28M13 rev2 from Amersham High quality sequence stop: 208. Location/Qualifiers 1..441 /organism="Mus musculus" /strain="C57BL/6J" /db_xref="taxon:10090" /clone="IMAGE:455332" /clone_lib="Soares mouse placenta 4NBP13.5 14.5" /sex="unknown" /tissue_type="Placenta" /dev_stage="adult" /lab_host="DH10B" /note="Organ: Placenta; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site.1: Not I; Site.2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer 5', TGTTACCAATCTGAAGTGGAGCGCGCGCAATTTTTTTTTTTTTTTTTTTT T 3.1; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT733 vector. Library went through one round of normalization, and was constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT	81 a 142 c 124 g 94 t
ORIGIN	
Alignment Scores:	
Pred. No.:	2,51e-32
Score:	44.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	4.00%
DB:	9 Matches: 441 Conservative: 44 Mismatches: 0 Indels: 0 Gaps: 0
US-09-397-967-16 (1-1099) x AA020387 (1-441)	
QY 1032 MetMetG1yProGluAR9G1G1yProPheLucySArgLeuAGluLeuLeuAGlu 1051	
Db 80 ATGATGGGGCGCTGACCGTGAAGAGACCCCGCTCTGCGCCTCTGGAGCTGTCGACAG 139	
QY 1052 G1yARgAgluEuproPProPProThCysPProThrcgluAgluAGluLeuMeG1leu 1071	
Db 140 GGGCGACGCCCTCCACACACCTCCACCTGCGCCGACGAGGTTCAGAGACTCATGCACTG 199	
QY 1072 CysTrPAlaPac 1075	
Db 200 TGCTGGGGGGCCC 211	
RESULT 30	

LOCUS	BE623895	501 bp	mRNA	linear	EST 24-AUG-2000
DEFINITION	Uti4d09_v1 Soares.thymus_2NBMT Mus musculus cDNA clone IMAGE:337857 5' similar to gb:L3768 Mus musculus (MOUSE);, mRNA sequence.				
ACCESSION	BE623895				
VERSION	BE623895				
KEYWORDS	EST.				
SOURCE	house mouse.				
ORGANISM	Mus musculus				
REFERENCE	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
AUTHORS	1 (bases 1 to 501)				
TITLE	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index				
JOURNAL	Unpublished (1997)				
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgaps-femail.nhl.gov This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information. NCI:1072021 Seq primer: -40RP from gldco High quality sequence stop: 469.				
FEATURES	Location/Qualifiers				
source	1..501				
	/organism="Mus musculus"				
	/strain="C57BL/6J"				
	/db_xref="taxon:10090"				
	/clone="IMAGE:337857"				
	/clone_lib="Soares_thymus_2NBMT"				
	/sex="male"				
	/tissue_type="Thymus"				
	/dev_stage="4 weeks"				
	/lab_host="DH10B"				
	/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker. Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTTACCAATCTGAAGTGAGCGCGGGTTTTTTTTTTTTTTTTTTT 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. RNA provided by Dr. Bertrand Jordan. Library went through two rounds of normalization, and was constructed by Bento Soares and M.Fatima Bonaldo."				
BASE COUNT	94 a 173 c 152 g 81 t				1 others
ORIGIN					
Alignment Scores:	2.92e-32	Length:	501		
Pred. No.:	44.00	Matches:	44		
Score:	100.00%	Conservative:	0		
Best local Similarity:	100.00%	Mismatches:	0		
Query Match:	4.00%	Indels:	0		
DB:	10	Gaps:	0		
US-09-397-967-16 (1-1099) x BE623895 (1-501)					
OY 1032 MetMetgLyProGLuaRgLuNgLyPProProlucySArGLEuLengluNeuleuAlagiu	1051				
Db 301 ATGATGGGGCCGTGAGGTAAAGACCCTGCCTCTGAGCTGCTGCCAGAG	360				
OY 1052 GLAARGALEuProPProPProFOrhrycSProTHgluValngludenuMetGInLeu	1071				
Db 361 GGGCGAGCGCTNCCACACTCCACACTCCCCACCGAGATTCAAGAGCTCATGCAGCTG	420				
OY 1072 CystRpAlapro	1075				
Db 421 TGCTGGGCCCCC	432				
RESULT 31					
AA015273 ,	597 bp	mRNA	linear	EST 21-JAN-1997	
LOCUS	AA015273				

DEFINITION	m30h10_r1 Soares mouse placenta 4NDMP13.5 14.5 Mus musculus cDNA clone IMAGE:444067 5' similar to gb:L22955 Mouse protein tyrosine kinase (MOUSE), mRNA sequence.					
ACCESSION	AA015273					
VERSION	AA015273.1 GI:1476305					
KEYWORDS	EST.					
SOURCE	house mouse.					
ORGANISM	Mus musculus					
REFERENCE	Mullayrola; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 597)					
AUTHORS	Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubnue,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Stepoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.					
TITLE	The Washu-HHMI Mouse EST Project					
JOURNAL	Unpublished (1996)					
COMMENT	Contact: Marra M/Mouse EST Project Washu-HHMI Mouse EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: mouseset@wustl.edu This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information. GI:269403 Putative full length read vector to vector length is 664 Seq primer:-28M13 rev2 from Amersham High quality sequence stop: 426. Location/Qualifiers 1..597 /organism="Mus musculus" /strain="C57BL/6J" /db_xref="taxon:10090" /clone="IMAGE:444067" /clone_lib="Soares mouse placenta 4NDMP13.5 14.5" /sex="unknown" /tissue_type="placenta" /dev_stage="adult" /lab_host="DH10B" /note="Organ: Placenta; Vector: pT73D-Pac (Pharmacia) with a modified polylinker. Site.1: Not I; Site.2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(drf) primer [5', TGTACCAATCTGAAGTGAGGACGCCCGGAAATTTTTTTTTTTTTTTTTC3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization, and was constructed by Bento Soares and M.Patima Bonaldo."					
BASE COUNT	107 a 199 c 175 g 115 t 1 others					
ORIGIN						
Alignment Scores:						
Pred. No.:	3	6e-32	Length:	597		
Score:	44.00		Matches:	44		
Percent Similarity:	100.00%		Conservative:	0		
Best Local Similarity:	100.00%		Mismatches:	0		
Query Match:	4.00%		Indels:	0		
DB:	9		Gaps:	0		
US-09-397-967-16 (1-1099) x AA015273 (1-597)						
OY	1032	MetMelGLyPRoGLuaRgLuGLyPProPrOLEuCysArgLeuLengluLeuEnuAgiu	1051			
Db	33	ATGATGGGGCCCTGAGCGGTGAAGACCCCCGCTCTGCCTCTGGAGCTGCGAGAG	92			
OY	1052	GLyARgArgeuRProPrOProTrInCySPyrThGlVAlGlnGluLeuNetsGlnu	1071			
Db	93	GggCGAgGcCTCCACCACCTCCACCCTGCCCCACCGAGAGTTAGAGAGCATGACGCTG	152			

DB:	12	Gaps:	0
US-09-397-967-16 (1-1099) x BF191625 (1-581)			
OY	989	TyrlaProglusleuserAspAnllePheSerArgGlnSerAspValTrpSerPhe	1008
Db	363	TATGCCCCGAGTCCCATCTGTGACACATCTTCTCCGCCAATCTGACGTGGAGCTTC	422
OY	1009	GIyAlValleuYrGluLeuPheMTrYrCYsAspLYsSerCYsSerProSerIaGlu	1028
Db	423	GGATGGTGTGTGTACGAGCTTTCACCTACTCGACACAGAGCTGCAGCCCATCCGCTAG	482
OY	1029	PhelEu 1030	
Db	483	TTTTT	
Db	483	TTCTCTG 488	
RESULT 33			
LOCUS	BF191625	554 bp	mRNA linear EST 02-NOV-2000
DEFINITION	239261 MARC 2P1G Sus scrofa cDNA 5', mRNA sequence.		
ACCESSION	BF191625		
VERSION	BF191625.1	GI:1074994	
KEYWORDS	EST.		
SOURCE	pig.		
ORGANISM	Sus scrofa		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.		
AUTHORS	1 (bases 1 to 554) Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E., Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Laegreid,W.W. and Keele,J.W.		
TITLE	Design and use of two pooled tissue normalized cDNA libraries for EST discovery in swine		
JOURNAL	Unpublished (2000)		
COMMENT	Contact: Smith TP USDA, ARS, US Meat Animal Research Center PO Box 166, Clay Center, NE 68933-0166, USA Tel: 402 762 4366 Fax: 402 762 4390 Email: smithemall@marc.usda.gov Single pass sequencing. Bases called and alt-trimmed with phred v0.980904.e. Vector identified by cross_match with the -m1nscore 18 and -m1match 12 options. PCR primers FORWARD: AGGAACAGCATGAGCAT BACKWARD: GTTTCCAGTCACGACG Plate: 67 row: E column: 8 Seq primer: ATTAGGTGACACTATGAG. Location/Qualifiers 1..554 /organism="Sus scrofa" /db_xref="taxon:9833" /clone_lib="MARC 2P1G" /tissue_type="pooled" /lab_host="DH10B" /note="Vector: PCMV SPOR6; Site_1: NOT1; Site_2: SalI; Library made from pooled tissue from testis, ovary, endometrium, hypothalamus, pituitary, and placenta."		
FEATURES	Source		
BASE COUNT	121 a 176 c 150 g 107 t		
ORIGIN			
Alignment Scores:			
Pred. No.:	2,85e-29	Length:	554
Score:	41.00	Matches:	41
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.73%	Indels:	0
DB:	12	Gaps:	0
US-09-397-967-16 (1-1099) x BF191625 (1-554)			
OY	860	GIaGlnArGAspRheGlnArgGlnIleGlnIleuLYsAlaLeuHISerAspPheIle	879

Db 432 CAGCAGAGGAGCTTCACGCGGAGATCCATCTTAAGCCCTCCACAGCAGCTTCATT 491

Qy 880 VallystYrArglyAlSerTyrglyPrgGlyAraGlnSerLeuArgLeuValMetGlu 899
 |||||||

Db 492 GTCAAGTACCGGGGTGTACACTATGAGCCAGCGCCAGAGTCTGGGGCTGTCTATGAA 551

Qy 900 Tyr 900
 |||

Db 552 TAT 554

RESULT 34
 AA920299 731 bp mRNA linear EST 20-APR-1998
 LOCUS AA920299
 DEFINITION v93002.r1 Soares.thymus.2NDMT Mus musculus cDNA clone
 IMAGE:1282802.5, similar to gb:LA0172 Mus musculus JAK3 gene,
 complete cds (MOUSE);, mRNA sequence.

ACCESSION AA920299
 VERSION AA920299.1 GI:3067078
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 731)
 Maier, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
 Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
 Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
 Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
 Waterston, R.
 The WashU-HM Mouse EST Project
 Unpublished (1996)
 CONTACT: Marra M/Mouse EST Project
 WASHU-HM Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:674602
 Seq primer: -28m13 rev2 ET from Amersham
 High quality sequence stop: 215.
 Location/Qualifiers
 1. 731
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="IMAGE:1282802"
 /clone_1ib="Soares.thymus.2NDMT"
 /sex="male"
 /tissue_type="thymus"
 /dev_stage="4 weeks"
 /lab_host="DH10B"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker. Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
 was primed with a Not I - oligo(dT) primer [5'
 TGTACCACTGACAGTGGAGCGCGGCGGTTTGTGTGTGTGTGTGTGTGTGT
 3']; double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified pT73 vector. RNA
 provided by Dr. Bertrand Jordan. Library went through two
 rounds of normalization, and was constructed by Bento
 Soares and M.Fatima Bonaldo."

BASE COUNT 167 a 209 c 214 g 141 t

ORIGIN

Alignment Scores:
 Pred. No.: 3.59e-27 Length: 731
 Score: 39.00 Matches: 39
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.55% Indels: 0

DB: 9 Gaps: 0

US-09-397-967-16 (1-1099) x AA920299 (1-731)

Qy 720 HisIleHrSerLeuGluProAlaLysLysLeuLysPheTyrgluAspGlnGlyGlnLeu 739
 |||||||

Db 187 CACATCACTGCTGAGAGCCGCCCAAAAGCTGAAGTTCTTGTAGGAGCAGCAGCAGCTG 246

Qy 740 ProAlaLeuLysTrpThgGluLeuIleThgGlnCysMetAlaTyAsp 758
 |||||||

Db 247 CCGGCTTCATAATGACAGACTGGGAGACTTATACACACATGCTGATGAT 303

RESULT 35
 AA023069 410 bp mRNA linear EST 21-JAN-1997
 LOCUS AA023069
 DEFINITION mh6d02.r1 Soares mouse placenta 4NDMP13.5 14.5 Mus musculus cDNA
 clone IMAGE:45907.5, similar to gb:L32955 Mouse protein tyrosine
 kinase (MOUSE);, mRNA sequence.

ACCESSION AA023069
 VERSION AA023069.1 GI:1487776
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 410)
 Maier, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
 Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
 Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
 Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
 Waterston, R.
 The WashU-HM Mouse EST Project
 Unpublished (1996)
 CONTACT: Marra M/Mouse EST Project
 WASHU-HM Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:272795
 Seq primer: -28M13 rev2 from Amersham
 High quality sequence stop: 299.
 Location/Qualifiers
 1. 410
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="IMAGE:45907"
 /clone_1ib="Soares mouse placenta 4NDMP13.5 14.5"
 /sex="unknown"
 /tissue_type="placenta"
 /dev_stage="adult"
 /lab_host="DH10B"
 /note="Organ: placenta; Vector: pT73D-Pac (Pharmacia)
 with a modified polylinker. Site 1: Not I; Site 2: Eco RI;
 1st strand cDNA was primed with a Not I - oligo(dT) primer
 [5'
 TGTACCACTGACAGTGGAGCGCGGCGGAAATTTTGTGTGTGTGTGTGTGTGT
 3']; double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified pT73 vector. Library
 went through one round of normalization, and was
 constructed by Bento Soares and M.Fatima Bonaldo."

BASE COUNT 67 a 152 c 128 g 62 t 1 others

ORIGIN

Alignment Scores:
 Pred. No.: 1.72e-26 Length: 410
 Score: 38.00 Matches: 38
 Percent Similarity: 100.00% Conservative: 0

US-09-397-967-16 (1-1099) x AA174246 (1-308)

OY 916 leuHSTHrAspArgLeuLeuLeuPheAlaTrpGlnIleCysGlyMetC1uTrpLeu 935
 |||||||
 Db 20 CTGCACACCGACCGCTACTGCTGTTCCTGGCAGATCTGCAGGGATGGAGTACCTG 79
 |||||||

OY 936 GlAlaAaGAGCyValAlaIsArgAspLeuAlaAlaAGsAlle 950
 |||||||
 Db 80 GGTGGCGCCGGTGGCTACCGTACCTGCTGGCGGCACATC 124
 |||||||

RESULT 37
 AM501263 231 bp mRNA linear EST 01-MAR-2000
 LOCUS
 DEFINITION
 UI-HF-BPDP-aiu-b-02-0-UI.r1 NIH_MGC_51 Homo sapiens cDNA clone
 IMAGE:3073082 5', mRNA sequence.
 AM501263
 ACCESSION
 VERSION
 KEYWORDS
 EST.
 SOURCE
 ORGANISM
 Homo sapiens
 human.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 NIH-MGC http://mgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: rgabs-remail.nih.gov
 Eco RI site shown at the beginning of the sequence.
 Tissue procurement: Louis M. Staudt, M.D., Ph.D.
 CDNA Library Preparation: M.B. Soares Lab
 CDNA Library Arrayed by: M.B. Soares Lab
 DNA Sequencing by: M.B. Soares Lab
 clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 www.bio.llnl.gov/bbrp/image/image.html
 Seq primer: M13 forward.

FEATURES
 source
 Location/Qualifiers
 1..231
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:3073082"
 /clone_1id="NIH_MGC_51"
 /tissue_type="lymph"
 /cell_type="geminal center B cells"
 /cell_line="MGC85"
 /lab_host="DH10B (L71)"
 /note="Vector: pRT73-Pac; Site_1: NotI; Site_2: Eco RI;
 Constructed from size fractionated cytoplasmic mRNA
 (4.4-7.4kb). Directionally cloned. Cells provided by
 Louis M. Staudt, Ph.D. Library preparation by Maria de
 Fatima Bonaldo, Ph.D. and M. Bento Soares, Ph.D."

BASE COUNT 35 a 90 c 53 g 52 t 1 others
 ORIGIN

Alignment Scores:
 Pred. No.: 7.18e-23 Length: 231
 Score: 34.00 Matches: 34
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.09% Indels: 0
 DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x AM501263 (1-231)

OY 994 leuSerAspAsnIlePheSerArgGlnSerAspValTrpSerPheGlyValValLeuTyr 1013
 |||||||
 Db 10 CTCTGGACACATCTTCTCTGCGCAGTCAGACCTTGGACCTTGGGGGTGCTCTGT 69
 |||||||

OY 1014 GluLeuPheThrTyrCysAspPlySerCysSerProSerAla 1027
 |||||||
 Db 70 GAGCTCTTCACCTACTCTGCGACAAAGGTGGAGCCCTCGGCC 111
 |||||||

RESULT 38
AM501582 254 bp mRNA linear EST 01-MAR-2000
LOCUS
DEFINITION UT-HF-BPDP-a-jg-b-03-0-UI.r1 NIH_MGC_51 Homo sapiens cDNA clone
IMAGE:3074236 5', mRNA sequence.
ACCESSION AM501582
VERSION AM501582.1 GI:7115248
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE NIH-MGC <http://mgc.nci.nih.gov/>
1 (bases 1 to 254)
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgapbs-remail.nih.gov
Eco RI site shown at the beginning of the sequence.
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
CDNA Library Preparation: M.B. Soares Lab
CDNA Library Arrayed by: M.B. Soares Lab
DNA Sequencing by: M.B. Soares Lab
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bdrip/image/image.html
Seq primer: M13 forward.

FEATURES
Source Location/Qualifiers
1..254
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3074236"
/clone_lib="NIH_MGC_51"
/tissue_type="lymph"
/cell_type="germinal center B cells"
/lab_host="DH10B (LTR)"
/note="Vector: pT73-Pac; Site_1: NotI; Site_2: Eco RI;
Constructed from size fractionated cytoplasmic mRNA
(4.4-7.4kb). Directionally cloned. Cells provided by
Louis M. Staudt, Ph.D. Library preparation by Maria de
Fatima Bonaldo, Ph.D. and M. Bento Soares, Ph.D."

BASE COUNT 37 a 97 c 64 g 55 t 1 others
ORIGIN

Alignment Scores:
Pred. NO.: 8.04e-23 Length: 254
Score: 34.00 Matches: 34
Percent Similarity: 100.00% Conservative: 0
Best local Similarity: 100.00% Mismatches: 0
Query Match: 3.09% Indels: 0
DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x AM501582 (1-254)

OY 994 LeuSerAspAsnIlePheSerArgInserAspValTTPSerPheGlyValValLeuTyr 1013
|||||
DB 10 CTCCTGGACACATCTTCTCTCGCAGTCAGACGCTGAGGCTCGGCTCTCTGAC 69
|||||

OY 1014 GluLeuPheThrTyrCysAspIlySerGlySerProSerAla 1027
|||||
DB 70 GAGCTCTTCACCTACTGCGACAAAGCTGACGCCCTCGGCC 111
|||||

RESULT 39
AM502902 286 bp mRNA linear EST 01-MAR-2000
LOCUS
DEFINITION UT-HF-BPDP-a-aiw-a-01-0-UI.r1 NIH_MGC_51 Homo sapiens cDNA clone
IMAGE:3073056 5', mRNA sequence.
ACCESSION AM502902
VERSION AM502902.1 GI:7117812
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE NIH-MGC <http://mgc.nci.nih.gov/>
1 (bases 1 to 286)
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgapbs-remail.nih.gov
Eco RI site shown at the beginning of the sequence.
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
CDNA Library Preparation: M.B. Soares Lab
CDNA Library Arrayed by: M.B. Soares Lab
DNA Sequencing by: M.B. Soares Lab
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bdrip/image/image.html
Seq primer: M13 forward.

FEATURES
Source Location/Qualifiers
1..286
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3073056"
/clone_lib="NIH_MGC_51"
/tissue_type="lymph"
/cell_type="germinal center B cells"
/lab_host="DH10B (LTR)"
/note="Vector: pT73-Pac; Site_1: NotI; Site_2: Eco RI;
Constructed from size fractionated cytoplasmic mRNA
(4.4-7.4kb). Directionally cloned. Cells provided by
Louis M. Staudt, Ph.D. Library preparation by Maria de
Fatima Bonaldo, Ph.D. and M. Bento Soares, Ph.D."

BASE COUNT 43 a 110 c 68 g 65 t
ORIGIN

Alignment Scores:
Pred. NO.: 9.26e-23 Length: 286
Score: 34.00 Matches: 34
Percent Similarity: 100.00% Conservative: 0
Best local Similarity: 100.00% Mismatches: 0
Query Match: 3.09% Indels: 0
DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x AM502902 (1-286)

OY 994 LeuSerAspAsnIlePheSerArgInserAspValTTPSerPheGlyValValLeuTyr 1013
|||||
DB 10 CTCCTGGACACATCTTCTCTCGCAGTCAGACGCTGAGGCTCGGCTCTCTGAC 69
|||||

OY 1014 GluLeuPheThrTyrCysAspIlySerGlySerProSerAla 1027
|||||
DB 70 GAGCTCTTCACCTACTGCGACAAAGCTGACGCCCTCGGCC 111
|||||

RESULT 40
AM501452 437 bp mRNA linear EST 01-MAR-2000
LOCUS
DEFINITION UT-HF-BPDP-a-jc-g-10-0-UI.r1 NIH_MGC_51 Homo sapiens cDNA clone
IMAGE:3074106 5', mRNA sequence.
ACCESSION AM501452
VERSION AM501452.1 GI:7114998
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE NIH-MGC <http://mgc.nci.nih.gov/>
1 (bases 1 to 437)
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgapbs-remail.nih.gov
Eco RI site shown at the beginning of the sequence.

Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
CDNA Library Preparation: M.B. Soares Lab
CDNA Library Arrayed by: M.B. Soares Lab
DNA Sequencing by: M.B. Soares Lab
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/dbrrp/image/image.html
Seq primer: M13 forward.

FEATURES
source

Location/Qualifiers
1. 437
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="3074106"
/clone_lib="NIH_MGC_51"
/tissue_type="lymph"
/cell_type="germinal center B cells"
/lab_host="DH10B (LT1)"
/note="Vector: p7773-Pac; Site_1: NotI; Site_2: Eco RI;
Constructed from size fractionated cytoplasmic mRNA
(4.4-7.4kb). Directionally cloned. Cells provided by
Louis M. Staudt, Ph.D. Library preparation by Maria de
Fatima Bonaldo, Ph.D. and M. Bento Soares, Ph.D."

BASE COUNT 63 a 180 c 105 g 89 t

ORIGIN

Alignment Scores:

Pred. No.: 1.53e-22 Length: 437
Score: 34.00 Matches: 34
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.09% Indels: 0
DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x AW501452 (1-437)

OY 994 LeuSerAspAsnIlePheSerArgInserAspValTrpSerPheGlyValValLeuTyr 1013
|||||
DB 10 CTCCTCGACACATCTCTCTCCAGTCAGAGCTGTGGGGCTGCTGTAC 69
|||||

OY 1014 GluLeuPheThrTyrCysAspLysSerCysSerProSerAla 1027
|||||
DB 70 GACCTCTTACCTACTGCGACAAAGCTGAGCCCTCGGCC 111
|||||

RESULT 41 451 bp mRNA linear EST 01-MAR-2000
AW500984
LOCUS UI-HF-BPOp-aly-f-11-0-UI-r1 NIH_MGC_51 Homo sapiens cDNA clone
DEFINITION IMAGE:3073676 5', mRNA sequence.
ACCESSION AW500984
VERSION AW500984.1 GI:7114087
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 451)
NIH-MGC http://mhc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Eco RI site shown at the beginning of the sequence.
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
CDNA Library Preparation: M.B. Soares Lab
CDNA Library Arrayed by: M.B. Soares Lab
DNA Sequencing by: M.B. Soares Lab
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/dbrrp/image/image.html
Seq primer: M13 forward.

FEATURES
source

Location/Qualifiers
1. 451

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="3073676"
/clone_lib="NIH_MGC_51"
/tissue_type="lymph"
/cell_type="germinal center B cells"
/lab_host="DH10B (LT1)"
/note="Vector: p7773-Pac; Site_1: NotI; Site_2: Eco RI;
Constructed from size fractionated cytoplasmic mRNA
(4.4-7.4kb). Directionally cloned. Cells provided by
Louis M. Staudt, Ph.D. Library preparation by Maria de
Fatima Bonaldo, Ph.D. and M. Bento Soares, Ph.D."

BASE COUNT 63 a 186 c 106 g 94 t 2 others

ORIGIN

Alignment Scores:

Pred. No.: 1.59e-22 Length: 451
Score: 34.00 Matches: 34
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.09% Indels: 0
DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x AW500984 (1-451)

OY 994 LeuSerAspAsnIlePheSerArgInserAspValTrpSerPheGlyValValLeuTyr 1013
|||||
DB 10 CTCCTCGACACATCTCTCTCCAGTCAGAGCTGTGGGGCTGCTGTAC 69
|||||

OY 1014 GluLeuPheThrTyrCysAspLysSerCysSerProSerAla 1027
|||||
DB 70 GACCTCTTACCTACTGCGACAAAGCTGAGCCCTCGGCC 111
|||||

RESULT 42 462 bp mRNA linear EST 16-FEB-2000
AW408832
LOCUS UI-HF-BMO-sdx-h-03-0-UI-r1 NIH_MGC_38 Homo sapiens cDNA clone
DEFINITION IMAGE:3063389 5', mRNA sequence.
ACCESSION AW408832
VERSION AW408832.1 GI:6927889
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 462)
NIH-MGC http://mhc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Eco RI site shown at the beginning of the sequence.
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
CDNA Library Preparation: M.B. Soares Lab
CDNA Library Arrayed by: M.B. Soares Lab
DNA Sequencing by: M.B. Soares Lab
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/dbrrp/image/image.html
Seq primer: M13 forward.

FEATURES
source

Location/Qualifiers
1. 462
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="3063389"
/clone_lib="NIH_MGC_38"
/tissue_type="lymph"
/cell_type="germinal center B cells"
/lab_host="DH10B (LT1)"
/note="Vector: p7773-Pac; Site_1: NotI; Site_2: Eco RI;
Constructed from size fractionated cytoplasmic mRNA

(2.5-3.5kb). Directionally cloned. Cells provided by Louis M. Staedt, Ph.D. Library preparation by Maria de Fatima Bonaldo, Ph.D. and M. Bento Soares, Ph.D."

BASE COUNT 86 a 156 c 131 g 89 t

ALIGNMENT SCORES:

Pred. No.: 1.64e-22 Length: 462
Score: 34.00 Matches: 34
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.09% Indels: 0
DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x AM408832 (1-462)

QY 938 ArgArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGluSerGluAlaHis 957

DB 265 CGCGGCTGCGTGCACGCGGACCTGCGCGCCGAAACATCTCTGAGAGCGAGGACACAC 324

QY 958 ValIysIleAlaAspPheGlyIleuAlaIysIleuProIeu 971

DB 325 GTCAAGATCGCTGACTTGGCTAGCTAAGCTGCTGCGCTT 366

RESULT 43

BF546077 478 bp mRNA linear EST 11-DEC-2000

LOCUS UI-R-C2P-rq-h-07-0-UI.r1 UI-R-C2P Rattus norvegicus cDNA clone

ACCESSION BF546077

VERSION BF546077.1 GI:11637184

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

COMMENT

1 (bases 1 to 478)
Bonaldo, M.F., Lennon, G., and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene discovery
Genome Res. 6 (9), 791-806 (1996)
Contact: Soares, MB
97044477
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: msoares@blue.weeg.uiowa.edu
cDNA Library Preparation: M.B. Soares Lab Clone distribution:
clones will be available through Research Genetics (www.resgen.com)
This clone is also available through the I.M.A.G.E. Consortium at
LNL (info@image.llnl.gov). IMAGE ID= 1790156
Seq primer: M13 Forward.

FEATURES

source

Location/Qualifiers
1..478
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-C2P-rq-h-07-0-UI"
/clone_id="UI-R-C2P"
/dev_stage="adult"
/lab_host="DHI0B (Life Technologies)"
/note="Vector: pUT73D-Pac (Pharmacia) with a modified polylinker. Site_1: Not I; Site_2: Eco RI; The UI-R-C2P library is a subtracted library derived from the UI-R-C1 library, which is a subtracted library derived from the UI-R-C0 library. The UI-R-C0 library consisted of a mixture of individually tagged normalized libraries constructed from rat placenta, adult lung, brain, liver, kidney, heart, spleen, ovary, muscle, 8, 12 and 18-day embryo. The tag is a string of 3-5 nucleotides present

between the Not I site and the oligo-dr track which allows identification of the library of origin of a clone within the mixture. The subtracted library (UI-R-C2P) was constructed as follows: PCR amplified cDNA inserts from UI-R-C1 clones from which 3' ESTs had been derived was used as a driver in a hybridization with the UI-R-C1 library in the form of single-stranded circles. The remaining single-stranded circles (subtracted library) was purified by hydroxypatite column chromatography, converted to double-stranded circles and electroporated into DH10B bacteria (Life Technologies) to generate the UI-R-C2P library. This procedure has been previously described (Bonaldo, Lennon and Soares, Genome Research 6: 791-806, 1996)"

BASE COUNT 83 a 167 c 128 g 99 t 1 others

ORIGIN

ALIGNMENT SCORES:

Pred. No.: 1.47e-19 Length: 478
Score: 31.00 Matches: 31
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.82% Indels: 0
DB: 12 Gaps: 0

US-09-397-967-16 (1-1099) x BF546077 (1-478)

QY 988 TrpIyrAlaProGluSerLeuSerAspAsnIlePheSerArgIleAspValTrpSer 1007

DB 9 TGTGATGCGCCCTGAGTCTGATCTGACACATCTTCTCCGCAATCTGACGTGTGAGC 68

QY 1008 PheGlyValIleuTyrgIleuPheThrTyr 1018

DB 69 TTGCGGCTGCTGTGATGAGCTCTTACCTAC 101

RESULT 44

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AM462965 483 bp mRNA linear EST 24-FEB-2000
BP230011A10C12 Soares normalized bovine placenta Bos taurus cDNA
clone BP230011A10C12 5', mRNA sequence.
AM462965
AM462965.1 GI:7033133
EST.
COW.
Bos taurus
Eukaryota; Metazoa; Chordata; Cranialta; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
Bovidae; Bovinae; Bos.
1 (bases 1 to 483)
Lewin, H.A., Soares, M.B., Rebeiz, M., Pardinas, J., Liu, L. and Larson, J.H.
Bovine ESTs
Unpublished (2000)
Contact: Lewin, H. A.
W. M. Keck Center for Comparative and Functional Genomics
University of Illinois at Urbana-Champaign
340 Edward R. Madigan Laboratory, 1201 W. Gregory Dr., Urbana, IL 61801, USA
Tel: 217 333 5998
Fax: 217 244 5617
Email: h-lewin@uiuc.edu
Funding for cattle EST sequencing was provided by the USDA National Research Initiative, Animal Genome Resource Grant AG 99-3205-6534 to H. A. Lewin and J. E. Womack. Base Calling/Quality Scores: PHRED from Washington University Genome Center. Vector Trimmi g:
Cross-match from Washington University Genome Center PHRAP suite.
Sequences submitted are vector free and at least 200 bp in length.
PCR Primers
FORWARD: TTAATACGACTCACTATAGG
BACKWARD: ATATACCTCCTCACTAAG
Insert length: 483 Std Error: 0.00
Plate: BP230011A10 row: C column: 12
Seq primer: AGCGATATACATTTTCACACAGGA

FEATURES High quality sequence stop: 483.
Location/Qualifiers

source

1. 483

/organism="Bos taurus"

/db_xref="taxon:9913"

/clone="BP23011A10C12"

/clone_lib="Soares normalized bovine placenta"

/sex="female"

/lab_host="DH10B"

/note="Organ: placenta; Vector: pT7T3pac; Site_1: ECORT; Site_2: NOTI; The cDNA library was contributed by the

Soares laboratory and it was constructed and normalized

as described by Bonaldo, M.F., Lennon, G. and Soares,

M.B. (1996), Genome Research 6(9): 791-806. "

BASE COUNT 83 a 185 c 121 g 94 t

ORIGIN

Alignment Scores:

Pred. No.: 1,42e-18

Score: 30.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 2,73%

DB: 10

US-09-397-967-16 (1-1099) x AW462965 (1-483)

QY 989 TTYAALPGLGSLERLEUSERASPAANLLEPHERARGINSEAPVALTPSERPHE 1008

DB 117 TATGCCCAAGATCCCTCTCGACACATCTTCTCGCGCCAGTCGGATGTCGAGCTTC 176

QY 1009 GLYVALVALLEUTYRGULEUPHERHYR 1018

DB 177 GGGGTGCTCCTGTACGAGCTCTCACCTAC 206

RESULT 45

AJ399251

LOCUS AJ399251 763 bp mRNA linear EST 25-JAN-2001

DEFINITION AJ399251 dkfz426 Gallus gallus cDNA clone 9d23r1, mRNA sequence.

ACCESSION AJ399251

VERSION AJ399251.1 GI:7134035

KEYWORDS EST.

SOURCE chicken.

ORGANISM Gallus gallus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Archosauria; Aves; Neognathae; Galliformes; Phasianidae;

Phasianinae; Gallus.

1 (bases 1 to 763)

Abdulkhmanov, I., Lodygin, D., Gerolt, P., Arakawa, H., Law, A., Plachy

J., Korn, B. and Buerstedde, J.M.

A large database of chicken bursal ESTs as a resource for the

analysis of vertebrate gene function

Genome Res. 10 (12), 2062-2069 (2000)

JOURNAL 20568495

MEDLINE Contact: Buerstedde JM

COMMENT Cellular Immunology

Heinrich-Pette-Institute

Martinistr. 52, 20251 Hamburg, Germany

Email: URL: http://genetics.hpi.uni-hamburg.de/dt40est.html.

FEATURES Location/Qualifiers

source 1. 763

/organism="Gallus gallus"

/strain="CB"

/db_xref="taxon:9031"

/clone="9d23r1"

/clone_lib="dkfz426"

/tissue_type="Bursa of Fabricius"

/cell_type="bursal lymphocyte"

/dev_stage="2-3 weeks old"

BASE COUNT 146 a 234 c 255 g 121 t 7 others

ORIGIN

Alignment Scores:

Pred. No.: 2,45e-18

Score: 30.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 2,73%

DB: 9

US-09-397-967-16 (1-1099) x AJ399251 (1-763)

QY 813 PHEGLGULARGHISLEULYSTYRIIESERLLEULGLYSGLYASNPHEGLYSERVAL 832

DB 254 TTGAGAGAGAGGACCACTCAAGTACATCTGCTGGCAAGGCACTTTGGAGCGTG 313

QY 833 GLULCYSARGTYRASPPOLEUGLYASP 842

DB 314 GAGCTGTCCGCTACGACCCGCTGGTGAC 343

RESULT 46

AJ442008

LOCUS AJ442008 769 bp mRNA linear EST 19-APR-2002

DEFINITION AJ442008 dkfz426 Gallus gallus cDNA clone 17j3r1, mRNA sequence.

ACCESSION AJ442008

VERSION AJ442008.1 GI:20209229

KEYWORDS EST.

SOURCE chicken.

ORGANISM Gallus gallus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Archosauria; Aves; Neognathae; Galliformes; Phasianidae;

Phasianinae; Gallus.

1 (bases 1 to 769)

Buerstedde, J.M.

Gallus gallus bursal lymphocyte EST

Unpublished (2002)

CONTACT: Buerstedde JM

Cellular Immunology

Heinrich-Pette-Institute

Martinistr. 52, 20251 Hamburg, Germany

Email: URL: http://genetics.hpi.uni-hamburg.de/dt40est.html.

FEATURES Location/Qualifiers

source 1. 769

/organism="Gallus gallus"

/strain="CB"

/db_xref="taxon:9031"

/clone="17j3r1"

/clone_lib="dkfz426"

/tissue_type="Bursa of Fabricius"

/cell_type="bursal lymphocyte"

/dev_stage="2-3 weeks old"

BASE COUNT 150 a 222 c 262 g 134 t 1 others

ORIGIN

Alignment Scores:

Pred. No.: 2,47e-18

Score: 30.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 2,73%

DB: 9

US-09-397-967-16 (1-1099) x AJ442008 (1-769)

QY 813 PHEGLGULARGHISLEULYSTYRIIESERLLEULGLYSGLYASNPHEGLYSERVAL 832

DB 127 TTGAGAGAGAGGACCACTCAAGTACATCTGCTGGCAAGGCACTTTGGAGCGTG 186

QY 833 GLULCYSARGTYRASPPOLEUGLYASP 842

DB 187 GAGCTGTCCGCTACGACCCGCTGGTGAC 216

RESULT 47

AM503697

LOCUS AM503697 411 bp mRNA linear EST 01-MAR-2000

DEFINITION UT-HF-BNO-akv-c-08-0-UI.r1 NIH_MGC_50 Homo sapiens cDNA clone

http://image.jnl.gov
Plate: LAM8531 row: f column: 14
High quality sequence stop: 532.
Location/Qualifiers
1.550

FEATURES

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone_id="IMAGE:3489805"
/clone_1lb="NCI_CGAP_Mam5"
/tissue_type="tumor, gross tissue"
/dev_stage="7 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies. Investigators providing samples: Iohar Hennighausen/Robin Humphreys, NIH"

BASE COUNT 97 a 153 c 158 g 142 t
ORIGIN

Alignment Scores:

Pred. No.: 1.43e-15 Length: 550
Score: 27.00 Matches: 27
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.46% Indels: 0
DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x BE309453 (1-550)

OY 111 GtyleuglunrCysHlsarGpHeGlyLeuArjySaspHeurHsrAlaileuAsp 130
|||||
Db 369 GGGCTGGAGACATCTCACCGCTTGGCTGGCAGAGATTGACCACTGCATCTTGC 428
|||||
OY 131 LeuHsVallleuglunHsleu 137
|||||
Db 429 TTACATGTATTAGAACATCTC 449

RESULT 50
BF191620 526 bp mRNA linear EST 02-NOV-2000
LOCUS 239255 MARC 2P1G Sus scrofa cDNA 5', mRNA sequence.
DEFINITION BF191620
ACCESSION BF191620.1 GI:11074989
VERSION
KEYWORDS EST.
SOURCE
ORGANISM Sus scrofa

REFERENCE Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi: Mammalia: Eutheria: Cetartiodactyla: Suina: Suidae; Sus.
AUTHORS Fahnenkrug, S.C., Freking, B.A., Rohrer, G.A., Smith, T.P.L., Casas, E., Stone, R.T., Heaton, M.P., Grose, W.M., Bennett, G.A., Laegreid, W.W., and Keele, J.W.

TITLE Design and use of two pooled tissue normalized cDNA libraries for EST discovery in swine
JOURNAL Unpublished (2000)
COMMENT Contact: Smith TPL

USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt_trimmed with phred v0.980904.e. Vector identified by cross_match with the -mnscore 18 and -mismatch 12 options.

FEATURES

PCR primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTCAGTCACGACG
Plate: 67 row: D column: 8
Seq primer: ATTACGTCACACTATG.
Location/Qualifiers
1.526

/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone_id="MARC 2P1G"
/tissue_type="pooled"
/lab_host="DH10B"
/note="Vector: pCMV SPORT6; Site_1: NotI; Site_2: SalI; Library made from pooled tissue from testis, ovary, endometrium, hypothalamus, pituitary, and placenta."

BASE COUNT 115 a 170 c 141 g 100 t
ORIGIN

Alignment Scores:

Pred. No.: 1.24e-13 Length: 526
Score: 25.00 Matches: 25
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.27% Indels: 0
DB: 12 Gaps: 0

US-09-397-967-16 (1-1099) x BF191620 (1-526)

OY 860 GtnglnArGAspHeGlnArGlnleGlnleuAlaileuHsSerAspHele 879
|||||
Db 432 CGCAGAGGAGACTTCCAGCGGAGATCCAGATCTTAAGCCCTCCACGCGACTTCA 491
|||||
OY 880 ValLySTyTArGly 884
|||||
Db 492 GTCAAGTACCGGGGG 506

RESULT 51
BM088287 564 bp mRNA linear EST 19-NOV-2001
LOCUS 501721 MARC 2BOV Bos taurus cDNA 5', mRNA sequence.
DEFINITION BM088287
ACCESSION BM088287.1 GI:16998915
VERSION
KEYWORDS EST.
SOURCE
ORGANISM Bos taurus

REFERENCE Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi: Mammalia: Eutheria: Cetartiodactyla: Ruminantia: Pecora: Bovidae; Bovidae; Bovine; Bos.
AUTHORS Smith, T.P.L., Grose, W.M., Freking, B.A., Roberts, A.J., Stone, R.T., Casas, E., Wray, J.E., White, J., Cho, J., Fahnenkrug, S.C., Bennett, G.L., Heaton, M.P., Laegreid, W.W., Rohrer, G.A., Chitko-McKown, C.G., Pertea, G., Holt, I., Karaymicheva, S., Liang, F., Quackenbush, J., and Keele, J.W.

TITLE Sequence evaluation of four pooled-tissue normalized bovine cDNA libraries and construction of a gene index for cattle
JOURNAL Genome Res. 11 (4), 626-630 (2001)
MEDLINE 21180013
COMMENT Contact: Smith TPL

USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt_trimmed with phred v0.980904.e. Vector identified by cross_match with the -mnscore 18 and -mismatch 12 options.

PCR primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTCAGTCACGACG
Plate: 140 row: C column: 18
Seq primer: ATTACGTCACACTATG.
Location/Qualifiers
1.564

FEATURES

/organism="Bos taurus"
/db_xref="taxon:9913"
/clone_id="MARC 2BOV"
/tissue_type="pooled"
/lab_host="DH10B"
/note="Vector: pCMV SPORT6; Site_1: NotI; Site_2: SalI;

Library made from pooled tissue from testis, thymus, semitendinosus muscle, longissimus muscle, pancreas, adrenal, and endometrium."

BASE COUNT 116 a 176 c 161 g 111 t

ALIGNMENT SCORES:

Pred. No.: 1.34e-13 Length: 564
Score: 25.00 Matches: 25
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.27% Indels: 0
DB: 13 Gaps: 0

US-09-397-967-16 (1-1099) x BM088287 (1-564)

OY 670 SerProthValleuSerLeuGlutMetLeuThrAspArgIleProTTPValAlaProGlu 689
|||||
DB 3 ACCCCACTGTGCTACAGCTGCTCAGTGCACAGAGATCCCTGGTGGCCCTGAA 62

OY 690 CysLeuGlnGluAla 694
|||||

DB 63 TGTCTCCAGAGAGGCC 77

RESULT 52

BO709825 871 bp mRNA linear EST 16-JUL-2002
LOCUS BO709825
DEFINITION AGENCOURT.8474895 NIH_MGC_113 Homo sapiens CDNA clone IMAGE:6301440
5', mRNA sequence.

ACCESSION BO709825.1 GI:21848724

KEYWORDS

EST.

SOURCE

human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

ORGANISM

Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE

1 (bases 1 to 871)
NIH-MGC <http://mgs.nci.nih.gov/>.

AUTHORS

National Institutes of Health, Mammalian Gene Collection (MGC)

TITLE

Unpublished (1999)

JOURNAL

Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov

COMMENT

Tissue Procurement: Dr. Mark Watson
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

FEATURES

DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LICM2517 row: f column: 01
High quality sequence stop: 595.
Location/Qualifiers

1..871

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:6301440"

/clone_id="NIH_MGC_113"

/lab_host="DH10B (phage-resistant)"

/note="Organ made by oligo-df priming. Site 1: XhoI; Site 2:
EcoRI; CDNA made by oligo-df priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGGACGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: This is a
NIH-MGC Library."

BASE COUNT 149 a 300 c 264 g 158 t

ORIGIN

Alignment Scores:

Pred. No.: 2.25e-13 Length: 871

Score: 25.00 Matches: 25

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.27% Indels: 0
DB: 14 Gaps: 0

US-09-397-967-16 (1-1099) x BO709825 (1-871)

OY 323 lLeleGluAlaGluPheProGlyLeuProGlyAlaLeuSerPheValAlaLeuValAsp 342
|||||

DB 328 ATTGTAGAGCGCGAGTTCCTCCAGGCGCTCCGAGGCGCTGTGCTGCGGGCGCTGGAG 387

OY 343 GlyTyrPheArgLeu 347
|||||

DB 388 GGCTACTTCCGGCTG 402

RESULT 53

BR791980 371 bp mRNA linear EST 16-NOV-2001
LOCUS BR791980
DEFINITION BR791980 RIKEN full-length enriched, brain CRT-1443 BC3H1 CDNA Mus
musculus cDNA clone GA30139C08.3, mRNA sequence.

ACCESSION BR791980.1 GI:16960744

KEYWORDS

EST.

SOURCE

house mouse.

ORGANISM

Mus musculus

REFERENCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 371)

Akimura,T., Arahata,T., Carninci,P., Furuno,M., Hanagaki,T., Ishii

Hayatsu,N., Hiramoto,K., Hirooka,T., Hirozane,T., Imclan,K., Ishii

Y., Ito,M., Kawai,D., Kojima,Y., Konno,H., Kouda,M., Matsuyama,T.,

Nakamura,M., Nishii,K., Nomura,K., Numasaki,R., Okazaki,Y., Okido,T.,

Saito,R., Sakai,C., Sakai,K., Sakazume,N., Sasaki,D., Sato,K.,

Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa

,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T.,

Watanahiki,A., Yasunishi,A., Morimatsu,M. and Hayashizaki,Y.,

RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al.

2001)

Unpublished (2001)

Contact: Yoshitake Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center(GSC) Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-res@sc.riken.go.jp,

URL:<http://genome-gsc.riken.go.jp/>

Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh

,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.,

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new

genes. Genome Res. 10 (10), 1617-1630 (2000)

wagui,K., Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,

Watanahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura

,S., Kawai,U., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and

Hayashizaki,Y.,

RIKEN integrated sequence analysis (RISA) system-384 format

sequencing pipeline with 384 multicapillary sequencer. Genome Res.

10 (11), 1757-1771 (2000)

Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara

,Y. and Hayashizaki,Y.,

Computer-based methods for the mouse full-length cDNA

encyclopedia: real-time sequence clustering for construction of a

nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

Please visit our web site (<http://genome-gsc.riken.go.jp>) for

further details.

e mouse tissues.

FEATURES

Location/Qualifiers

1..371

/organism="Mus musculus"

/strain="C3H"

/db_xref="taxon:10090"

/clone="G430139C08"

```

/dev_stage="10 months, virgin"
/lab_host="DH10g"
/note="Organ: mammary; Vector: pcwv-SPORT6; site_1: SalI;
site_2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"
BASE COUNT      85 a      113 c      79 g      61 t      1 others
ORIGIN

Alignment Scores:
Pred. No.:      6,66e-12      Length:      339
Score:      23,00      Matches:      23
Percent Similarity:      100.00%      Conservative:      0
Best Local Similarity:      100.00%      Mismatches:      0
Query Match:      2.09%      Indels:      0
DB:      10      Gaps:      0

US-09-397-967-16 (1-1099) x AM230832 (1-339)
Qy      1053      ArgArgLeuProPProPProPProhCysProhGluValGlnGluLeuMetGlnLeuGys      1072
|||||
Db      3      CGAGCGCTCCACACACCTCCACCTCCACCGAGGTTGAGAGCTCATGACGCTGTC      62
|||||
Qy      1073      TrpAlaPro      1075
|||||
Db      63      TGGGCGGCC      71

RESULT 55
AM505513      540 bp      mRNA      linear      EST 02-MAR-20000
LOCUS      UI-HF-BNO-amb-g-09-0-UI.r1 NIH_MGC-50 Homo sapiens cDNA clone
ACCESSION      IMAGE:3081784 5', mRNA sequence.
VERSION      AM505513
KEYWORDS      AM505513.1 GI:7143180
SOURCE      EST.
ORGANISM      human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 540)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Eco RI site shown at the beginning of the sequence.
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
cDNA Library Preparation: M.B. Soares Lab
cDNA Library Arrayed by: M.B. Soares Lab
DNA Sequencing by: M.B. Soares Lab
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LINL at:
www-bio.lnlnl.gov/bdrp/image/image.html
Seq primer: M13 Forward.
Location/Qualifiers
1..540
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3081784"
/clone_1lb="NIH_MGC-50"
/tissue_type="lymph"
/cell_type="germinal center B cells"
/lab_host="DH10B (LTI)"
/note="Vector: pTrf3-Pac; Site_1: NotI; site_2: Eco RI;
Constructed from size fractionated cytoplasmic mRNA
(3.5-4kb). Directionally cloned. Cells provided by
Louis M. Staudt, Ph.D. Library preparation by Maria de
Fatima Bonaldo, Ph.D. and M. Bento Soares, Ph.D."
BASE COUNT      116 a      164 c      151 g      108 t      1 others
ORIGIN

```

Pred. No.: 1.1e-10 Length: 540
 Score: 22.00 Matches: 22
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 2.00% Indels: 0
 DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x AW505513 (1-540)

OY 824 LeuGLYfsglyAsnPhgelySerValGluLeuCyArGfyrAspProLeuGLYfsglyAspAsn 843
 |||||||
 DB 365 CTGGCAAGGGCACTTGGCAGCGTGGAGCTGTGCCGTATGACCGCTAGCGACAAAT 424

OY 844 ThrGLY 845
 |||||||
 DB 425 ACAGCT 430

RESULT 56
 AW785501

LOCUS 472 bp mRNA linear EST 09-JUL-2000
 DEFINITION 116874 MARC 1P1G Sus scrofa cDNA 5', mRNA sequence.
 ACCESSION AW785501 GI:7842349

VERSION
 KEYWORDS
 SOURCE
 ORGANISM

Sus scrofa
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
 1 (bases 1 to 472)
 Fahnenkrug,S.C., Freking,B.A., Rohrer,G.A., Shultz,T.P.L., Casas,E.,
 Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Laegreid,W.W.,
 and Keefe,J.W.
 Design and use of two pooled tissue normalized cDNA libraries for
 EST discovery in swine
 Unpublished (2000)

JOURNAL
 COMMENT

Contact: Smith TPL
 USDA, ARS, US Meat Animal Research Center
 PO Box 166, Clay Center, NE 68933-0166, USA
 Tel: 402 762 4366
 Fax: 402 762 4390

Email: smith@email.marc.usda.gov
 Simple pass sequencing. Bases called and alt-trimmed with phred
 v0.980904.e. Vector identified by cross-match with the -mismoscore 18
 and -mismatch 12 options.

PCR Primers
 FORWARD: AGGAACAGCATATGACCAT
 BACKWARD: GTTTCACAGTACGACG
 Plate: 52 row: D column: 2
 Seq primer: ATTATGATGACATATAG.

FEATURES
 source
 1. 472
 Location/Qualifiers

/organism="Sus scrofa"
 /db_xref="taxon:9823"
 /clone_lib="MARC 1P1G"
 /tissue_type="pooled"
 /lab_host="DH10B"
 /note="Vector: PCMV SPORT6; Site_1: NCI; Site_2: Sali;
 library made from pooled tissue from day 11, 13, 15, 20,
 and 30 embryos."

BASE COUNT 117 a 128 c 134 g 93 t
 ORIGIN

Alignment Scores:
 Pred. No.: 8.95e-10 Length: 472
 Score: 21.00 Matches: 21
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.91% Indels: 0
 DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x AW785501 (1-472)

OY 578 HistLeuValLeuHisGLYfValCysMetValGlyAspSerIleMetValGlnGluPhe 597

DB 322 CATCTGTTGCTTCATGCGCTATGCTGAGACAGCATCATGTCGAGAAATTC 381
 |||||||
 OY 598 Val 598
 |||||
 DB 382 GTA 384

RESULT 57
 BB590115

LOCUS 258 bp mRNA linear EST 30-NOV-2000
 DEFINITION BB590115 RIKEN full-length enriched, 0 day neonate thymus Mus
 musculus cDNA clone AA30103J21 5', mRNA sequence.
 BB590115

ACCESSION
 BB590115 GI:11486659

VERSION
 KEYWORDS
 SOURCE
 ORGANISM

REFERENCE
 AUTHORS

1 (bases 1 to 258)
 Aizawa,K., Akahira,S., Akimura,T., Arai,A., Arakawa,T., Carninci,P.,
 Hanagaki,T., Hayatsu,N., Hiraoka,T., Hirozane,T., Hodojima,Y.,
 Imocani,K., Ishii,Y., Itoh,M., Izawa,M., Kawai,J., Kojima,Y., Konno
 H., Kusakabe,M., Matsuyama,T., Miyazaki,A., Nakamura,M., Nishi,K.,
 Nomura,K., Numazaki,K., Okazaki,Y., Okido,T., Owa,C., Sakai,C.,
 Sakai,K., Sasaki,D., Sato,K., Shibata,K., Shibata,Y., Shingawa,A.,
 Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa,A., Takahashi,F., Tanaka
 T., Toyota,T., Watahiki,A., Yamamura,T., Yasunishi,A., Yoshida,K.,
 Yoshiki,A., Muramatsu,M., and Hayashizaki,Y.
 RIKEN Mouse ESTs (Aizawa,K. et al. 2000)
 Unpublished (2000)

JOURNAL
 COMMENT

Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center(GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suenho-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216

Email: genome-resgsc.riken.go.jp/
 URL: http://genome.gsc.riken.go.jp/
 Carninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoka,S., Sakai
 N., Okazaki,Y., Muramatsu,M., and Hayashizaki,Y.
 Thermotabilization and thermostabilization of thermolabile enzymes by
 trehalose and its application for the synthesis of full length
 cDNA. Proc Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)

Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M., Okazaki
 Y., and Hayashizaki,Y.
 Automated filtration-based high-throughput plasmid preparation
 system. Genome Res. 9 (5), 463-470 (1999)

High-efficiency full-length cDNA cloning. Methods Enzymol. 303,
 19-44 (1999)
 Please visit our web site (http://genome.rtc.riken.go.jp) for
 further details.

FEATURES
 source

1. 258
 Location/Qualifiers
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone_lib="AA30103J21"
 /clone_lib="RIKEN full-length enriched, 0 day neonate
 thymus"

/tissue_type="thymus"
 /dev_stage="0 day neonate"
 /lab_host="DH10B"

/note="Site_1: Sali; Site_2: BamHI; cDNA library was
 prepared and sequenced in Mouse Genome Encyclopedia
 Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in
 RIKEN. Division of Experimental Animal Research in Riken
 contributed to prepare mouse tissues. 1st strand cDNA was
 primed with a primer 15'
 GAGAGAGAGAGATCCAGAGCTCTTTTCTTTTCTTTTCTTNN 3', cDNA was

prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 459.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGATTCTGATTAATTAATTCCTCCCTCC 3']. cDNA was cleaved with XhoI and BamHI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLX I."

BASE COUNT 46 a 69 c 74 g 69 t

ORIGIN

Alignment Scores:

Pred. No.:	Score:	Percent Similarity:	Best Local Similarity:	Query Match:	DB:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
3	786-07	18.00	100.00%	1.64%	10	258	18	0	0	0	0

US-09-397-967-16 (1-1099) x BB590115 (1-258)

Oy 17 SerIeuSerSerSerGluAlaAlaAlaHisValIeuLeuProPargIy 34
 Db 77 AGCCTCTATCCGACAGAGAGAGCCCTGCAATGCTCTCTCCCGGGA 130

RESULT 58
 A1651109 291 bp mRNA linear EST 16-DEC-1999
 LOCUS wa97d05.x1 NCI CGAP GC6 Homo sapiens cDNA IMAGE:2304105 3'
 DEFINITION similar to SW:JAK3_HUMAN P52333 TYROSINE-PROTEIN KINASE JAK3 ;,
 mRNA sequence.

ACCESSION A1651109
 VERSION A1651109.1 GI:4735088
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 291)
 AUTHORS NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cga@pds-remail.nih.gov
 Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.

CDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CCGAP clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at:
 www-bio.illn.gov/dbrrp/image/image.html
 Insert length: 328 Std Error: 0.00
 Seq primer: -40up from Gibco.

FEATURES

Location/Qualifiers

1..291

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:2304105"

/clone_lib="NCI CGAP GC6"

/tissue_type="pooled germ cell tumors"

/lab_host="DH10B"

/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker. Site_1: Not I; Site_2: Eco RI; Plasmid DNA from the normalized library NCI CGAP GC4 was prepared, and ss circles were made in vitro. Following HAT purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (clonids 1257096-1258631, 1469064-1470983, and 1475592-1476743).

Subtraction by Bento Soares and M. Fatima Bonaldo. "

BASE COUNT 65 a 87 c 96 g 43 t

ORIGIN

Alignment Scores:

Pred. No.:	Score:	Percent Similarity:	Best Local Similarity:	Query Match:	DB:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
4	366-07	18.00	100.00%	1.64%	9	291	18	0	0	1	0

US-09-397-967-16 (1-1099) x A1651109 (1-291)

Oy 1 MetAlaProSerSerGluThrProLeuIleProGlnArgSerCysSerIeu 18
 Db 265 ATGGCACCTCCAGTAGAGAGAGAGCCCGTGAATCCCTGACGCTTACGAGCTTC 212

RESULT 59
 BM032194 319 bp mRNA linear EST 05-NOV-2001
 LOCUS 497815 MARC 2BOV Bos taurus cDNA 5', mRNA sequence.
 DEFINITION BM032194
 ACCESSION BM032194.1 GI:16745764
 VERSION BM032194.1 GI:16745764
 KEYWORDS EST.
 SOURCE cow.
 ORGANISM Bos taurus

REFERENCE 1 (bases 1 to 319)
 AUTHORS Smith,T.P.L., Grose,W.M., Freking,B.A., Roberts,A.J., Stone,R.T., Casas,E., Wray,J.E., White,J., Cho,J., Fahnenkrug,S.C., Bennett,G.L., Heaton,M.P., Laegreid,W.M., Rohrer,G.A., Chilton-McKown,C.G., Pertea,G., Holt,I., Karanmycheva,S., Liang,F., Quackenbush,J. and Keeler,J.W.
 TITLE Sequence evaluation of four pooled-tissue normalized bovine cDNA libraries and construction of a gene index for cattle
 JOURNAL Genome Res. 11 (4), 626-630 (2001)
 MEDLINE 21180013
 COMMENT Contact: Smith TPL
 USDA, ARS, US Meat Animal Research Center
 PO Box 166, Clay Center, NE 68933-0166, USA
 Tel: 402 762 4366
 Fax: 402 762 4390
 Email: smith@email.marc.usda.gov
 Single pass sequencing. Bases called and alt trimmed with phred v0.980904.e. Vector identified by cross-match with the -m1nsc0re 18 and -m1nmatch 12 options.

PCR Primers
 FORWARD: AGGAACAGCTATGACCAT
 BACKWARD: GTTTCCTCCAGTCACGACG
 Plate: 134 row: H column: 24
 Seq primer: ATTGAGGACACATATAC.
 Location/Qualifiers

FEATURES

1..319

/organism="Bos taurus"

/db_xref="taxon:9913"

/clone_lib="MARC 2BOV"

/tissue_type="pooled"

/lab_host="DH10B"

/note="Vector: pCMV SPORT6; Site_1: NotI; Site_2: SalI; Library made from pooled tissue from testis, thymus, semitendonsus muscle, longissimus muscle, pancreas, adrenal, and endometrium."

BASE COUNT 67 a 88 c 86 g 78 t

ORIGIN

Alignment Scores:

Pred. No.:	Score:	Percent Similarity:	Best Local Similarity:	Mismatches:
4	866-07	18.00	100.00%	0

Length: 319
 Matches: 18
 Conservative: 0
 Mismatches: 0

Query Match: 1.64% Indels: 0
DB: 13 Gaps: 0

US-09-397-967-16 (1-1099) x BM032194 (1-319)

QY 559 AsncyswtcgtgusrcphcgluAlaAlaLaserLeuMetSerGlnValSerTyr 576
|||||
DB 112 AATTGCATGAGCATTCCTCTGGAAGACGAGCTTGATGACCAAGTCTCATAC 59

RESULT 60
BM146700 373 bp mRNA linear EST 30-NOV-2001
LOCUS TCAAP1E7063 Pediatric acute myelogenous leukemia cell (FAB M1)
DEFINITION Baylor-HSC Project-TCAA Homo sapiens CDNA clone TCAAP7063, mRNA
sequence.
ACCESSION BM146700
VERSION BM146700.1 GI:17165865
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 373)
AUTHORS Wei, Y., Tsang, Y.T.M., Mei, G., Ku, J.M., Ali-Osman, F.R. Jr.,
Gunsarane, P.H., Muzny, D., Bouck, J., Gibbs, R.A. and Margolin, J.F.
TITLE Pediatric leukemia CDNA Sequencing Project (2001)
JOURNAL Unpublished (2001)
COMMENT Contact: Dr. Judith F. Margolin
Texas Children's Cancer Center and Human Genome Sequencing Center
at Baylor College of Medicine
1102 Bates, MC3-3320 Houston, TX 77030, USA
Tel: 832-824-4536
Fax: 832-825-4038
Email: clones@ccc.org
Seq primer: M13 primer.
Location/Qualifiers
1. 373
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="TCAAP7063"
/clone_1bp="Pediatric acute myelogenous leukemia cell (FAB
M1) Baylor-HSC Project-TCAA"
/sex="male"
/tissue_type="leukocytes"
/cell_type="myeloid cell"
/dex_stage="myeloid 6 years"
/lab_host="DH10B"
/note="Vector: lambda PSB; Site 1: BamHI; Site 2: EcoRI;
First strand cDNA was primed with an anchored
XhoI-oligo(dT) primer [5'GGAGAGCTGACGCGCGGAGAGAG(17)VN
3'; V-A,C,G; N-A,C,G,T] and then dg tailed. Second strand
was primed with a BamHI-dc primer
[5'AGAGAGCTCGATCCGCGCGCGCATATATATAT(C) 3'].
Double-stranded cDNA was then digested with BamHI and XhoI
and directionally cloned into the BamHI and SalI sites of
lambda PSB vector. Library was constructed by Wei Yu at RIKEN
of Japan (Carninci P., Westover A., Nishiyama Y., Ohsumi T.,
Itoh M., Nagaoka S., Sasaki, Okazaki Y., Muramatsu M.,
Schneider C., Hayashizaki Y. High efficiency selection of
full-length cDNA by improved biotinylated cap trapper.,
DNA Res 4: 1, 61-6, Feb 28, 1997)."

BASE COUNT 51 a 138 c 101 g 83 t
ORIGIN

Alignment Scores: 5.85e-07 Length: 373
Pred. No.: 18.00 Matches: 18
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 1.64% Gaps: 0
DB: 13

US-09-397-967-16 (1-1099) x BM146700 (1-373)

QY 1 MetAlaProSerGluGluThrProLeuIleProGlnArgSerCysSerLeu 18
|||||
DB 63 ATGGACCTCCAAAGTGAAGAGACGCGCTTGATCCCTCAGCGTTCATGACGCTC 116

RESULT 61
AM504672 413 bp mRNA linear EST 202-MAR-2000
LOCUS UT-HF-BNO-all-d-09-0-UT.r1 NIH-MGC_50 Homo sapiens CDNA clone
DEFINITION IMAGE:3080104 5', mRNA sequence.
ACCESSION AM504672
VERSION AM504672.1 GI:7142339
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 413)
AUTHORS NIH-MGC http://mgs.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Eco RI site shown at the beginning of the sequence.
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
CDNA Library Preparation: M.B. Soares Lab
CDNA Library Arrayed by: M.B. Soares Lab
DNA Sequencing by: M.B. Soares Lab
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bdrp/image/image.html
The following repetitive elements were found in this cDNA
sequence:
121-152. >GC rich flow complexity
Seq primer: M13 Forward
Location/Qualifiers
1. 413
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3080104"
/clone_1bp="NIH-MGC_50"
/tissue_type="lymph"
/cell_type="germinal center B cells"
/lab_host="DH10B (LTI)"
/note="Vector: p77r3-Pac; Site 1: NotI; Site 2: Eco RI;
Constructed from size fractionated cytoplasmic mRNA
(3.5-4.4kb). Directionally cloned. Cells provided by
Louis M. Staudt, Ph.D. Library preparation by Maria de
Fatima Bonaldo, Ph.D. and M. Bento Soares, Ph.D."

BASE COUNT 63 a 142 c 114 g 94 t
ORIGIN

Alignment Scores: 6.61e-07 Length: 413
Pred. No.: 18.00 Matches: 18
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 1.64% Gaps: 0
DB: 10

US-09-397-967-16 (1-1099) x AM504672 (1-413)

QY 1 MetaIaProSerGluGluThrProLeuIleProGlnArgSerCysSerLeu 18
|||||
DB 32 ATGGACCTCCAAAGTGAAGAGACGCGCTTGATCCCTCAGCGTTCATGACGCTC 85

RESULT 62
AO192794 481 bp DNA linear GSS 04-NOV-1998
LOCUS HS-2251_B2_P07_MR CIT Approved Human Genomic Sperm Library D Homo
DEFINITION sapiens genomic clone Plate-2251 Col-14 Row-L, DNA sequence.

ACCESSION AQ192794
 VERSION AQ192794.1 GI:3591416
 KEYWORDS GSS.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 481)
 AUTHORS Mahalax,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
 Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
 Hood,L.
 TITLE Sequence-tagged connectors: A sequence approach to mapping and
 scanning the human genome
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
 MEDLINE 99380589
 COMMENT Contact: Mahalax GG, Wallace JC, Hood L
 High Throughput Sequencing Center
 University of Washington
 401 Queen Anne Avenue North, Seattle, WA 98109, USA
 Tel: (206) 616-3618
 Fax: (206) 616-3887
 Email: jwallace@u.washington.edu
 Sequence Tagged Connector
 Plate: 2251 row: L column: 14
 Class: BAC ends
 High quality sequence stop: 481.
 Location/Qualifiers
 1..481
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_plate="2251 Col-14 Row-L"
 /clone_lib="CIT Approved Human Genomic Sperm Library D"
 /sex="male"
 /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
 E-Coli DH10B"
 BASE COUNT 101 a 115 c 192 g 72 t 1 others
 ORIGIN
 Alignment Scores:
 Pred. No.: 7.92e-07 Length: 481
 Score: 18.00 Matches: 18
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.64% Indels: 0
 DB: 17 Gaps: 0
 US-09-397-967-16 (1-1099) x AQ192794 (1-481)
 QY 943 ArgaspleuAlalaarganlleuValguSerGluAlaHisVallystle 960
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 432 CGCGACCTGGCCGCCGGAACAATCCTCGTGAGAGCGACACGCTCAAGATC 379
 RESULT 63 781 bp mRNA linear EST 16-OCT-2001
 BI906136
 LOCUS 603062556r1 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5211901 5',
 DEFINITION mRNA sequence.
 ACCESSION BI906136
 VERSION BI906136.1 GI:16168788
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 781)
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-r@mail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL, at:
 http://image.llnl.gov
 Plate: LLML1531 row: p column: 14
 High quality sequence stop: 581.
 Location/Qualifiers
 1..781
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_image="5211901"
 /clone_lib="NIH_MGC_118"
 /tissue_type="Leukocyte"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV
 (destroyed); RNA source leukocytes from anonymous pool of
 non-activated adult donors. Library is oligo-dT primed
 and directionally cloned (EcoRV site is destroyed upon
 cloning). Average insert size 1.7 kb, insert size range
 1.2-3.3 kb. Library is normalized and enriched for
 full-length clones and was constructed by C. Gruber
 (Invitrogen). Research Genetics tracking code 027. Note:
 this is a NIH_MGC Library."
 BASE COUNT 242 a 207 c 184 g 148 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1.41e-06 Length: 781
 Score: 18.00 Matches: 18
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.64% Indels: 0
 DB: 13 Gaps: 0
 US-09-397-967-16 (1-1099) x BI906136 (1-781)
 QY 559 AsncysmetGluSerPheLeuGluAlaAlaSerLeuMetSerGlnValSerTyr 576
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 480 AACTGCATGTGAGTCAATCTCGGAGAGCGACGCTGATGAGCCAACTGTCTGAC 533
 RESULT 64 814 bp mRNA linear EST 16-OCT-2001
 BI911448
 LOCUS 603063362r1 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5212575 5',
 DEFINITION mRNA sequence.
 ACCESSION BI911448
 VERSION BI911448.1 GI:16175189
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 814)
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-r@mail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLML1533 row: 1 column: 16
 High quality sequence stop: 5
 High quality sequence stop: 416.
 Location/Qualifiers
 1..814
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_image="5212575"

/clone_lib="NIH_MGC_118"
 /tissue_type="leukocyte"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV
 (destroyed); RNA source leukocytes from anonymous pool of
 non-activated adult donors. Library is oligo-dT primed
 and directionally cloned (EcoRV site is destroyed upon
 cloning). Average insert size 1.7 kb, insert size range
 1.2-3.3 kb. Library is normalized and enriched for
 full-length clones and was constructed by C. Gruber
 (Invitrogen). Research Genetics tracking code 027. Note:
 this is a NIH_MGC Library."
 BASE COUNT 168 a 258 c 237 g 151 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1.48e-06 Length: 814
 Score: 18.00 Matches: 18
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.64% Indels: 0
 DB: 13 Gaps: 0
 US-09-397-967-16 (1-1099) x BI911448 (1-814)
 Oy 1 MetalapProSerGluGluThrProLeuIleProGlnArgSerCysSerIeu 18
 Db 90 ATGGACCTCCAGTAGAGAGACGCCCTGATCCCTCATGCGATGACGCTC 143
 RESULT 65
 LOCUS BF147824 313 bp mRNA linear EST 29-DEC-2000
 DEFINITION uy83e09.x1 NCI CGAP Mam5 Mus musculus cDNA clone IMAGE:366184 3'
 similar to SW:JAK3_RAT Q63272 TYROSINE-PROTEIN KINASE JAK3 ;, mRNA
 sequence:
 BF147824
 BF147824.1 GI:11029219
 EST.
 house mouse.
 Mus musculus
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 1 (bases 1 to 313)
 NCI CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 Other_ESTs: uy83e09.y1
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 image.lnl.gov/image/html/lresources.shtml
 MGI:1426952
 Seq primer: -40UP from GlbCO
 High quality sequence stop: 280.
 Location/Qualifiers
 1..313
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="IMAGE:366184"
 /clone_lib="NCI CGAP Mam5"
 /tissue_type="tumor, gross tissue"
 /dev_stage="7 months"
 /lab_host="DH10B"
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
 Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
 FEATURES
 source

```

Library constructed by Life Technologies. Investigators
providing samples: Lothar Hennighausen/Robin Humphreys,
NIH"

BASE COUNT      61 a      81 c      111 g      55 t      5 others
ORIGIN

Alignment Scores:
Pred. No.:      4,53e-06      Length:      313
Score:          17.00      Matches:      17
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      1.55%      Indels:      0
DB:              12      Gaps:      0

US-09-397-967-16 (1-1099) x BF147824 (1-313)      !

QY      1059      ProthRGSPProthGluValGlnGluMetGlnLeuGlyTPALAPro 1075
Db      286      CCCACTGCCCCACGAGGTTGAGGAGCTCATGACAGCTGTGCTGGCGGCC 236

RESULT 66
LOCUS      BG995133/c      426 bp      mRNA      linear      EST 13-JUN-2001
DEFINITION CMO-HIT295-160201-779-g06 HIT295 Homo sapiens cDNA, mRNA sequence.
ACCESSION      BG995133
VERSION      BG995133.1      GI:14399203
KEYWORDS
SOURCE      human.
ORGANISM      Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 426)
              Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
              Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
              Goldman,G.H., Carvalho,A.F., Matsumura,A., Bala,G.S., Simpson,D.H.,
              Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
              ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
              Simpson,A.J.
              Shotgun sequencing of the human transcriptome with ORF expressed
              sequence tags
              Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?l=CM0&t2=CM0-HIT295
160201-779-g06&t3=2001-02-16&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 10
High quality sequence stop: 426.
Location/Qualifiers
1. 426
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT1295"
/dev_stage="Adult"
/note="Organ: head,neck; Vector: puc18; Site:1; Smat:
Site-2; Smat: A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
FEATURES
Source
BASE COUNT      62 a      120 c      122 g      102 t
ORIGIN

```


Alignment Scores:

Pred. No.: 6.53e-06 Length: 426
 Score: 17.00 Matches: 17
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.55% Indels: 0
 DB: 13 Gaps: 0

US-09-397-967-16 (1-1099) x BG995133 (1-426)

QY 514 ApsSerleuGlutPhisGluAnleuGlyHisGlySerPheThrLysIle 530
 |||||||

DB 187 GACAGCCTGAGTGGATGAGACCTGGCCATGGCTTCACCAAGATT 137

RESULT 67 430 bp mRNA linear EST 09-MAR-1999
 A1472755 taa1e03.x1 NCI CGAP_Lym5 Homo sapiens cDNA clone IMAGE:2043964 3'
 LOCUS similar to SM:JAK3_HUMAN P52333 TYROSINE-PROTEIN KINASE JAK3 ;
 DEFINITION mRNA sequence.

ACCESSION A1472755

VERSION A1472755

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 1 (bases 1 to 430)
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)

JOURNAL Contact: Robert Strausberg, Ph.D.
 Email: cgapb-remail.nih.gov

COMMENT Tissue Procurement: Mark Raffeld, M.D.
 CDNA Library Preparation: Stratagene, Inc.
 CDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LMW at:
 www.bio.lnl.gov/bdrp/image/image.html

Trace considered overall poor quality
 Seg primer: -40UP from Glibco
 High quality sequence stop: 1.
 Location/Qualifiers
 1. .430
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2043964"
 /clone_lib="NCI-CGAP_Lym5"
 /issue_type="follicular lymphoma"
 /lab_host="SOLR (Stratagene, karyomycin resistant)"
 /note="Organ: lymph node; Vector: pBluescript SK-; Site_1:
 EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer:
 Oligo dT. Average insert size 1.2 kb. Non-amplified
 library. -5' adaptor sequence: 5' GAATTGGCAGGAG 3' -3'
 adaptor sequence: 5' CTCGACTTTT TTTT TTTT 3' "

FEATURES

SOURCE

BASE COUNT 95 a 135 c 97 g 103 t
 ORIGIN

Alignment Scores:

Pred. No.: 6.6e-06 Length: 430
 Score: 17.00 Matches: 17
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.55% Indels: 0
 DB: 9 Gaps: 0

US-09-397-967-16 (1-1099) x A1472755 (1-430)

OY 560 CysMetGluSerPheLeuGluAlaIleSerLeuMetSerGlnValSerTyr 576
 |||||||

DB 94 TGCATGAGTCAATCTCTGAGACGACGAGCTTGATGACCAAGTGTCTAC 44

RESULT 68

LOCUS A1598913/C

DEFINITION A1598913

ACCESSION EST250616

KEYWORDS Normalized rat embryo, Bento Soares Rattus sp. cDNA clone

SOURCE Rattus sp.

ORGANISM Rattus sp.

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 449)
 Lee,N.H., Gloder,A., Chandra,I., Mason,T.M., Quackenbush,J.,
 Kerlavage,A.R. and Adams,M.D.
 Rat Genome Project: Generation of a Rat EST (RST) Catalog & Rat
 Gene Index
 Unpublished (1998)
 Other_ESTs: TC82898
 Contact: Lee, NH
 The Institute for Genomic Research
 9712, Medical Center Drive, Rockville, MD 20850, USA
 Tel: (301)-838-3529
 Fax: (301)-838-0208
 Email: nhlee@tigr.org
 Seq primer: M13-21.

JOURNAL Location/Qualifiers

COMMENT 1. .449

ORGANISM="Rattus sp."

db_xref="taxon:10118"

/clone="R6MD161"

/clone_lib="Normalized rat embryo, Bento Soares"

/dev_stage="embryo 8, 12, 18 dpc"

/note="Vector: pT713Pac; Site_1: EcoRI; Site_2: NotI"

BASE COUNT 102 a 118 c 135 g 94 t
 ORIGIN

Alignment Scores:

Pred. No.: 6.95e-06 Length: 449
 Score: 17.00 Matches: 17
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.55% Indels: 0
 DB: 9 Gaps: 0

US-09-397-967-16 (1-1099) x A1598913 (1-449)

QY 159 GlnGlyGluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArg 175
 |||||||

DB 420 CAGGAGAGTTCCTGACCTAGCTGTCTGACCTGCCACAGATGGCTCGT 370

RESULT 69

LOCUS BB285471

DEFINITION BB285471

ACCESSION BB285471

KEYWORDS RIKEN full-length enriched 2 cells egg Mus musculus cDNA

SOURCE clone B020006C2.3, similar to I1656 Mouse protein-tyrosine kinase

ORGANISM (JAK2) mRNA, mRNA sequence.

REFERENCE BB285471

AUTHORS BB285471

1 (bases 1 to 239)
 Kono,H., Aizawa,K., Akahira,S., Akiyama,J., Arakawa,T., Carrinoci,
 P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N.,
 Hirozane,T., Hori,F., Ishii,Y., Ishikawa,J., Ishikawa,T., Itoh,M.,
 Izawa,M., Kadota,K., Kagawa,I., Kai,C., Kawai,J., Kikuchi,N.,
 Kiyosawa,H., Kojima,Y., Kondo,S., Koye,S., Kurihara,C., Kusakabe,M.

FEATURES	Location/Qualifiers
source	1. . 239

Alignment Scores:	
Pred. No.:	3,13e-05
Score:	16.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	1.46%
DB:	10
Length:	239
Matches:	16
Conservative:	0
Mismatches:	0
Indels:	0
Gaps:	0

```

FEATURES
source
Location/Qualifiers
1. .258

```

US-09-397-967-16 (1-1099) x BF67544 (1-258)

OY 1003 SerSpAltaTPSPePhGlyValValIleuArgIleuPheThrTyr 1018
| | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | |
DB 154 TCAGATGTTTGACCTTGGACTGCTTCGTATGAACCTTTCACATAC 107

RESULT 71

BB285470	304 bp	mRNA	linear	EST 09-JUL-2000
BB285470	RIKEN full-length enriched,	2 cells egg Mus musculus	cDNA	
clone BB20006C31	3' similar to J1656	Mouse protein-tyrosine kinase		
(Jtk2)	mRNA, mRNA sequence.			

DB: 9 Gaps: 0

US-09-397-967-16 (1-1099) x AA453345 (1-398)

QY 1003 SeraspValTPSerPheGlyValValLeuTyrgtuleuPhehrTyr 1018
 |||||||
 DB 207 TCAGATGTTTGAGCTTGAGAGTGTGTGTATGAACTTTCACATAC 254

RESULT 73
 T48592 406 bp mRNA linear EST 02-FEB-1995
 LOCUS ph6h6_19/1TV Outward Alu-primed hncDNA library Homo sapiens cDNA
 DEFINITION clone ph6h6_19/1TV, mRNA sequence.
 T48592
 ACCESSION T48592.1 GI:642792
 VERSION EST.
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 406)
 Vinogradova,T.V., Lebedev,Y.B., Kopantzev,E.P., Wagner,L.L., Volik
 S.V., Ermolaeva,O.D., Lavrenyeva,I., Monastyrskaya,G.S. and
 Sverdlov,E.D.
 Outward Alu-primed hncDNA library
 TITLE Unpublished (1995)
 JOURNAL Contact: Sverdlov ED
 COMMENT Structure and Function of Human Genes
 Shemyakin Institute of Bioorganic Chemistry
 16/10 Miklukho-Maklaya, Moscow, 117871, Russia
 Tel: 70953306529
 Fax: 70953306538
 Email: sverdlohungen.slobc.msk.su.

FEATURES
 Source location/Qualifiers
 1..406
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="ph6h6_19/1TV"
 /clone_lib="Outward Alu-primed hncDNA library"
 /note="Vector: pGEM-3Z; Site 1: EcoRI; Site 2: BamHI; The
 library was constructed as described in [Obradovic, D.,
 Borodin, A.M., Kopantzev, E.P., Wagner, L.L., Volik, S.V.,
 Ermolaeva, O.D., Lebedev, Y.B., Monastyrskaya, G.S.,
 Sverdlov, E.D. (1993) Bioorganicheskaya khimiya, 20,
 919-930]. This protocol is based on nested primer strategy
 using Alu- specific primers (ALN3 and TC-65) that direct
 the hncDNA synthesis outward of Alu repeats."

BASE COUNT 80 a 105 c 99 g 122 t

ORIGIN

Alignment Scores:
 Pred. No.: 5.88e-05 Length: 406
 Score: 16.00 Matches: 16
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.46% Indels: 0
 DB: 14 Gaps: 0

US-09-397-967-16 (1-1099) x T48592 (1-406)

QY 875 HisSeraspPheIleValIysTyraGlyValISerTygIleProGly 890
 |||||||
 DB 25 CACAGTGAATTCATGTCAGATGTCGTGTCACCTATGCCCGAGGT 72

RESULT 74
 BM147781 430 bp mRNA linear EST 30-NOV-2001
 LOCUS TCAAP1014273 Pediatric acute myelogenous leukemia cell (FAB M1)
 DEFINITION Baylor-HGSC project-TCAA Homo sapiens cDNA clone TCAAP1427, mRNA
 sequence.
 ACCESSION BM147781
 VERSION BM147781.1 GI:17167905
 KEYWORDS EST.

SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 430)
 Wei,Y., Tsang,Y.T.M., Mei,G., Ku,J.M., Ali-Osman,F.R., Jr.,
 Gunaratne,P.H., Munny,D., Bouck,J.G., Gibbs,R.A. and Margolin,J.F.
 Pediatric Leukemia cDNA Sequencing Project (2001)
 TITLE JOURNAL Unpublished (2001)
 COMMENT Contact: Dr. Judith F. Margolin
 Texas Children's Cancer Center and Human Genome Sequencing Center
 at Baylor College of Medicine
 1102 Bates, MC3-3320 Houston, TX 77030, USA
 Tel: 832-824-4536
 Fax: 832-825-4038
 Email: clones@txccc.org
 Seq primer: M13 primer.

FEATURES
 Source location/Qualifiers
 1..430
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="TCAAP1427"
 /clone_lib="pediatric acute myelogenous leukemia cell (FAB
 M1) Baylor-HGSC project-TCAA"
 /sex="male"
 /tissue="leukopheresis"
 /cell_type="myeloid cell"
 /dev_stage="pediatric 6 years"
 /lab_host="DH10B"
 /note="Vector: lambda PSB; Site 1: BamHI; Site 2: EcoRI;
 First strand cDNA was primed with an anchored
 XhoI-oligo(dT) primer [5'GGAGACTGTGACGGCCGACGAGAG(T)VN
 3'; V=A,C,G; N=A,C,G,T] and then dg tailed. Second strand
 was primed with a BamHI-dc primer
 [5'AGAGAGCTGACGACCGCGCGCAATATATAT(C) 3'].
 Double-stranded cDNA was then digested with BamHI and XhoI
 and directionally cloned into the BamHI and SalI sites of
 lambda PSB vector. Library went through one round of
 normalization. Library was constructed by Wei Yu at RIKEN
 of Japan (Carinaci P, Westover A, Nishiyama Y, Ohsumi T,
 Itoh M, Nagaoka S, Sasaki, Okazaki Y, Muramatsu M,
 Schneider C, Hayashizaki Y. High efficiency selection of
 full-length cDNA by improved biofinlyated cap trapper.,
 DNA Res 4: 1, 61-6, Feb 28, 1997)"]

BASE COUNT 59 a 149 c 126 g 96 t

ORIGIN

Alignment Scores:
 Pred. No.: 6.29e-05 Length: 430
 Score: 16.00 Matches: 16
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.46% Indels: 0
 DB: 13 Gaps: 0

US-09-397-967-16 (1-1099) x BM147781 (1-430)

QY 62 GlyIleLeuProValTyrrHisSerleuPheAlaLeuAlaInGluasp 77
 |||||||
 DB 248 GGCATCTGCGCTGTACACATCCCTCTTGTCTGGCAGCAGAGAC 295

RESULT 75
 BG364706 532 bp mRNA linear EST 08-MAR-2001
 LOCUS dc75a07.y1 NICHD XGC Embl Xenopus laevis cDNA clone IMAGE:3402900
 DEFINITION 5' similar to TR-019064 019064 JMK2. ;, mRNA sequence.
 ACCESSION BG364706
 VERSION BG364706.1 GI:13253803
 KEYWORDS EST.
 SOURCE African clawed frog.
 ORGANISM Xenopus laevis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;

FEATURES	high quality sequence stop: /53.
source	Location/Qualifiers
	1. .753
	/ordname="Mus musculus"

BASE COUNT 249 a 153 c 188 g 163 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 0.000122 Length: 753
 Score: 16.00 Matches: 16
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.46% Indels: 0
 DB: 12 Gaps: 0

US-09-397-967-16 (1-1099) x B6870416 (1-753)
 OY 1003 SeraspValTrpSerPheGlyValValLeuTyrGluLeuPheThrTyr 1018
 DB 397 TCAGATGCTGTGAGCTTGGAGTGTCTATACGAACTTTTCACATAC 444

RESULT 78
 B1153342 766 bp mRNA linear EST 05-JUL-2001
 LOCUS 602918614F1 NCI_CGAP_Lu29 Mus musculus cDNA IMAGE:5068946 5'
 DEFINITION mRNA sequence.
 B1153342
 B1153342 GI:14613343
 ACCESSION
 VERSION B1153342.1
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE NIH-MGC http://mgc.nci.nih.gov/
 AUTHORS 1 (bases 1 to 766)
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Gilbert Smith, Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLM11185 row: 1 column: 03
 High quality sequence stop: 760.

FEATURES
 source
 Location/Qualifiers
 1..766
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="IMAGE:5068946"
 /clone_1ib="NCI_CGAP_Lu29"
 /tissue_type="spontaneous tumor, metastatic to mammary."
 /lab_host="DHI0B"
 /stem_cell_origin=""
 /note="Organ: Lung; Vector: PCMV-SPORT6; Site: 1; Salt; Site: 2; NotI: Cloned unidirectionally. Primer: Oligo dt. Library constructed by Life Technologies. Investigator providing samples: Gilbert Smith, NIH"

BASE COUNT 254 a 145 c 194 g 173 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 0.000125 Length: 766
 Score: 16.00 Matches: 16

Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.46% Indels: 0
 DB: 13 Gaps: 0

US-09-397-967-16 (1-1099) x B1153342 (1-766)
 OY 1003 SeraspValTrpSerPheGlyValValLeuTyrGluLeuPheThrTyr 1018
 DB 444 TCAGATGCTGTGAGCTTGGAGTGTCTATACGAACTTTTCACATAC 491

RESULT 79
 B0899824 920 bp mRNA linear EST 16-AUG-2002
 LOCUS AGENCOURT_8755626 NIH_MGC_130 Mus musculus cDNA clone IMAGE:6334890
 DEFINITION 5', mRNA sequence.
 B0899824
 B0899824 GI:22291838
 ACCESSION
 VERSION B0899824.1
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE NIH-MGC http://mgc.nci.nih.gov/
 AUTHORS 1 (bases 1 to 920)
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Mark Macconochie, Ph.D. and Nancy L. Freeman, Ph.D.
 cDNA Library Preparation: ResGen, Invitrogen Corp
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLM13796 row: 9 column: 19
 High quality sequence stop: 663.

FEATURES
 source
 Location/Qualifiers
 1..920
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="IMAGE:6334890"
 /clone_1ib="NIH_MGC_130"
 /lab_host="DHI0B (phage-resistant)"
 /note="Organ: Oocytes; Vector: PCMV-SPORT6.1.ccd; Site: 1; EcoRV; Site: 2; NotI: Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.95 kb. Constructed by ResGen, Invitrogen Corp. Note: this is a NIH-MGC Library."

BASE COUNT 295 a 192 c 216 g 215 t 2 others
 ORIGIN
 Alignment Scores:
 Pred. No.: 0.000155 Length: 920
 Score: 16.00 Matches: 16
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.46% Indels: 0
 DB: 14 Gaps: 0

US-09-397-967-16 (1-1099) x B0899824 (1-920)

OY 1003 SeraspValTrpSerPheGlyValValLeuTyrGluLeuPheThrTyr 1018
 DB 288 TCAGATGCTGTGAGCTTGGAGTGTCTATACGAACTTTTCACATAC 335
 RESULT 80
 B0732700 945 bp mRNA linear EST 16-JUL-2002
 LOCUS AGENCOURT_8097616 NICHD XGC Emb4 Xenopus laevis cDNA clone
 DEFINITION IMAGE:5572666 5', mRNA sequence.

ACCESSION B0732700.1 GI:21871597
 VERSION EST.
 KEYWORDS African clawed frog.
 SOURCE ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
 Xenopodinae; Xenopus.
 1 (bases 1 to 945)
 NCI-CCAP <http://www.ncbi.nlm.nih.gov/nciccap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 CONTACT: Robert Strausberg, Ph.D.
 Email: cgaps-r@mail.nih.gov
 Tissue Procurement: Dr. Igor David
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: NCI-CCAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bbrp/image/image.html
 Plate: L1AM12318 row: h column: 11
 High quality sequence stop: 352.
 Location/Qualifiers
 1..945
 /organism="Xenopus laevis"
 /db_xref="taxon:8355"
 /clone="IMAGE:5572666"
 /clone_1lb="NICHD XGC Emb4"
 /dev_stage="embryo, stage 31-32"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: whole embryo; Vector: pCMV-SPORT6; Site: 1;
 Note: Site 2: Salt; Cloned unidirectionally. Primer: Oligo
 dt. Average insert size 2.1 kb. Constructed by Life
 Technologies. Note: This is a Xenopus Gene Collection (XGC
) library."
 BASE COUNT 315 a 197 c 217 g 215 t 1 others
 ORIGIN
 Alignment Scores:
 Pred. No.: 0.00016 Length: 945
 Score: 16.00 Matches: 16
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.46% Indels: 0
 DB: 14 Gaps: 0
 US-09-397-967-16 (1-1099) x B0732700 (1-945)
 QY 1003 SerAspValTrpSerPheGlyValValLeuTyGluLeuPheThyr 1018
 |||||||
 Db 304 TCAGATGCTGGAGCTTGGAGTTGCTCTATGAGTGTTCACATAC 351
 RESULT 81
 CDS042XQ 989 bp DNA linear GSS 18-MAY-2000
 LOCUS Tetraodon nigroviridis genome survey sequence PUC-ori end of clone
 DEFINITION 078f11 of library G from Tetraodon nigroviridis, genomic survey
 sequence.
 ACCESSION AL212087.1 GI:7994146
 VERSION AL212087.1
 KEYWORDS GSS; genome survey sequence.
 SOURCE Tetraodon nigroviridis.
 ORGANISM Tetraodon nigroviridis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
 Tetraodontidae; Tetraodon.
 1 (bases 1 to 989)
 Roest-Crollius, H., Jallion, O., Dasilva, C., Bouneau, L., Fisher, C.,
 Bernot, A., Fitzames, C., Wincker, P., Brottier, P., Quetier, F.,
 Saurin, W. and Weissenbach, J.

TITLE Human gene number estimate provided by genome wide analysis using
 JOURNAL Tetraodon nigroviridis DNA sequence
 REFERENCE Unpublished
 2 (bases 1 to 989)
 Roest-Crollius, H., Jallion, O., Dasilva, C., Fitzames, C., Fisher, C.,
 Bouneau, L., Billault, A., Quetier, F., Saurin, W., Bernot, A. and
 Weissenbach, J.
 TITLE Characterization and repeat analysis of the compact genome of the
 freshwater pufferfish Tetraodon nigroviridis
 JOURNAL Unpublished
 3 (bases 1 to 989)
 Roest-Crollius, H., Jallion, O., Dasilva, C., Bouneau, L., Fisher, C.,
 Bernot, A., Fitzames, C., Wincker, P., Brottier, P., Quetier, F.,
 Saurin, W. and Weissenbach, J.
 TITLE Human gene number estimate provided by genome wide analysis using
 JOURNAL Tetraodon nigroviridis DNA sequence
 REFERENCE Unpublished
 2 (bases 1 to 1001)
 Roest-Crollius, H., Jallion, O., Dasilva, C., Fitzames, C., Fisher, C.,
 Bouneau, L., Billault, A., Quetier, F., Saurin, W., Bernot, A. and
 Weissenbach, J.
 TITLE Characterization and repeat analysis of the compact genome of the
 freshwater pufferfish Tetraodon nigroviridis
 JOURNAL Unpublished
 3 (bases 1 to 1001)
 BASE COUNT 237 a 288 c 275 g 184 t 5 others
 ORIGIN
 Alignment Scores:
 Pred. No.: 0.000169 Length: 989
 Score: 16.00 Matches: 16
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.46% Indels: 0
 DB: 17 Gaps: 0
 US-09-397-967-16 (1-1099) x CDS042XQ (1-989)
 QY 1003 SerAspValTrpSerPheGlyValValLeuTyGluLeuPheThyr 1018
 |||||||
 Db 414 TCCGACCTTGGAGCTTGGCGCTGCTCTACGAGCTTTCACCTAC 461
 RESULT 82
 CDS038MW/C 1001 bp DNA linear GSS 15-MAY-2000
 LOCUS Tetraodon nigroviridis genome survey sequence PUC-ori end of clone
 DEFINITION 005j23 of library G from Tetraodon nigroviridis, genomic survey
 sequence.
 ACCESSION AL232817.1 GI:7891952
 VERSION AL232817.1
 KEYWORDS GSS; genome survey sequence.
 SOURCE Tetraodon nigroviridis.
 ORGANISM Tetraodon nigroviridis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
 Tetraodontidae; Tetraodon.
 1 (bases 1 to 1001)
 Roest-Crollius, H., Jallion, O., Dasilva, C., Bouneau, L., Fisher, C.,
 Bernot, A., Fitzames, C., Wincker, P., Brottier, P., Quetier, F.,
 Saurin, W. and Weissenbach, J.
 TITLE Human gene number estimate provided by genome wide analysis using
 JOURNAL Tetraodon nigroviridis DNA sequence
 REFERENCE Unpublished
 2 (bases 1 to 1001)
 Roest-Crollius, H., Jallion, O., Dasilva, C., Fitzames, C., Fisher, C.,
 Bouneau, L., Billault, A., Quetier, F., Saurin, W., Bernot, A. and
 Weissenbach, J.
 TITLE Characterization and repeat analysis of the compact genome of the
 freshwater pufferfish Tetraodon nigroviridis
 JOURNAL Unpublished
 3 (bases 1 to 1001)

AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (12-APR-2000)
COMMENT This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the tetraodon nigroviridis genome. For more information, please take a look at <http://www.genoscope.cns.fr/tetraodon>.

FEATURES
source
1. 1001
/organism="Tetraodon nigroviridis"
/db_xref="taxon:99883"
/clone_lib="G"
/note="Genoscope sequence ID : COBG005CE12XEL-end : PUC-ori"

BASE COUNT 209 a 265 c 274 g 249 t 4 others

ORIGIN

Alignment Scores:
Pred. No.: 0.000172 Length: 1001
Score: 16.00 Matches: 16
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.46% Indels: 0
DB: 17 Gaps: 0

US-09-397-967-16 (1-1099) x CNS038MM (1-1001)

OY 1003 Serpaspaltrpserpseglyvalvalleuylrgluenuepethrrtyr 1018
|||||
Db 273 TCCGACGTTTGAGCTTGCGGCTGCTCTACGAGCTCTTACCTTAC 226

RESULT 83
BF373713/ 253 bp mRNA linear EST 24-NOV-2000
LOCUS BF373713 PM4-SN0016-080600-005-b03 SN0016 Homo sapiens cDNA, mRNA sequence.
DEFINITION BF373713
VERSION BF373713.1 GI:11335829
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 253)
Nagai,M.A., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W.Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bata,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (<http://www.ludwig.org.br/scripts/gethtml2.pl?tl-PM4<2-PM4-SN0016-080600-005-b03<3-2000-06-08<4-1>)
Seq primer: puc 18 forward
High quality sequence start: 17
High quality sequence stop: 66.
Location/Qualifiers
1. 253
/organism="Homo sapiens"
/db_xref="taxon:9606"

FEATURES
source

/clone_lib="SN0016"
/dev_stage="Adult"
/note="Organ: stomach_normal; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORSTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 58 a 79 c 60 g 56 t

ORIGIN

Alignment Scores:
Pred. No.: 0.000319 Length: 253
Score: 15.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.36% Indels: 0
DB: 12 Gaps: 0

US-09-397-967-16 (1-1099) x BF373713 (1-253)

OY 600 LeuGLyAlaIleAspMetYrLeuArglySarGLyHisLeuVal 614
|||||
Db 240 CTGGGGCCATAGACATGTATCTGGAAACGTGCCACCTGGTG 196

RESULT 84
BH043390 754 bp DNA linear GSS 17-JUN-2001
LOCUS BH043390 RPCI-24-360L21.TJ RPCI-24 Mus musculus genomic clone RPCI-24-360L21
DEFINITION RPCI-24-360L21.TJ RPCI-24 Mus musculus genomic clone RPCI-24-360L21
ACCESSION BH043390
VERSION BH043390.1 GI:14825476
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 754)
Zhao,S., Nieman,W., Malek,J., Shatsman,S., Akincet,B., Levins,M., Tesgaye,G., Geer,K., Krol,M., Shvartsbeyn,A., Gebregeorgis,E., Russell,D., de Jong,P. and Fraser,C.M.
Mouse BAC End Sequences from Library RPCI-24
Unpublished (1999)
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the mouse BAC library RPCI-24. For BAC library availability, please contact Pieter de Jong (pdejong@tigr.org). Clones may be purchased from BACPAC Resources (<http://www.chori.org/bacpac/orderingframe.htm>). BAC end page: http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
Plate: 360 row: L column: 21
Seq primer: SP6
Class: BAC ends.
Location/Qualifiers
1. 754
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-24-360L21"
/clone_lib="RPCI-24"
/sex="Male"
/cell_type="Spleen/Brain"
/note="Vector: pTRABAC1; Site_1: BamHI; Site_2: BamHI; RPCI-24 Mouse BAC Library produced by Pieter de Jong. The library was cloned in the pTRABAC1 cloning vector at the BamHI sites using MboI partially digested male C57BL/6J DNA."

FEATURES
source

BASE COUNT 144 a 224 c 210 g 176 t
 ALIGNMENT SCORES:
 Pred. No.: 0.00117 Length: 754
 Score: 15.00 Matches: 15
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.36% Indels: 0
 DB: 17 Gaps: 0
 US-09-397-967-16 (1-1099) x BH043390 (1-754)
 QY 333 GUAAlaUSeRpheValaLeuValaSpGLYTYr-PheArgLeu 347
 ||||||||||||||||||||||||||||||||||||||||||||
 Db 584 GAGGCCCTGTCTTGTGGCCGTGCGATGCTATTTCGGCTTG 628
 RESULT 85
 LOCUS AM437524 281 bp mRNA linear EST 25-APR-2001
 DEFINITION 78773 MARC IBOV Bos taurus cDNA 5', mRNA sequence.
 ACCESSION AM437524
 VERSION AM437524.1 GI:6972830
 KEYWORDS EST.
 SOURCE cow.
 ORGANISM Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 Bovidae; Bovinae; Bos.
 1 (bases 1 to 281)
 Smith,T.P.L., Grose,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
 Caess,E., Wray,J.E., White,J., Cho,J., Fahnenkrug,S.C., Bennett,
 G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-Mckown,C.G.,
 Pertea,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,J. and
 Keeler,J.W.
 Sequence evaluation of four pooled-tissue normalized bovine cDNA
 libraries and construction of a gene index for cattle
 Genome Res. 11 (4), 626-630 (2001)
 JOURNAL 21180013
 COMMENT Contact: Smith TPL
 USDA, ARS, US Meat Animal Research Center
 PO Box 166, Clay Center, NE 68933-0166, USA
 Tel: 402 762 4366
 Fax: 402 762 4390
 Email: smitht@mail.marc.usda.gov
 Single pass sequencing. Bases called and trimmed with phred
 v0.980904.e. Vector identified by cross-match with the -minscore 20
 and -mismatch 12 options.
 PCR Primers
 FORWARD: AGGAACAGCATGACCAT
 BACKWARD: GTTTCAGTCACGACG
 Plate: 42 row: A column: 9
 Seq primer: ATTAGGTCACATATAG.
 FEATURES
 source
 Location/Qualifiers
 1..281
 /organism="Bos taurus"
 /db_xref="taxon:9913"
 /clone_lib="MARC IBOV"
 /tissue_type="pooled"
 /lab_host="DH10B"
 /note="Vector: PCMV SPORT6; site_1: NotI; site_2: SalI;
 library made from pooled tissue from lymph node, ovary,
 fat, hypothalamus, and pituitary."
 BASE COUNT 41 a 95 c 91 g 54 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 0.00345 Length: 281
 Score: 14.00 Matches: 14
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.27% Indels: 0
 DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x AM437524 (1-281)
 QY 334 AAlaUSeRpheValaLeuValaSpGLYTYr-PheArgLeu 347
 ||||||||||||||||||||||||||||||||||||||||||||
 Db 58 GCGTTGTCTTGTCTGCGCTGCTGATGCTACTTCCGGCTG 99
 RESULT 86
 LOCUS AM212845/c 296 bp mRNA linear EST 03-DEC-1999
 DEFINITION u068d02.x1 NCI CGAP Maml Mus musculus cDNA clone IMAGE:2647683 3',
 similar to gb:U40172 Mus musculus JAK3 gene, complete cds (MOUSE);,
 mRNA sequence.
 ACCESSION AM212845
 VERSION AM212845.1 GI:6518985
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 Tissue Procurement: Gilbert Smith, Ph.D.
 CDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Washington University Genome Sequencing Center
 clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 www-bio.llnl.gov/bdrp/image/image.html
 MGI:1028135
 Seq primer: -400P from gibco.
 FEATURES
 source
 Location/Qualifiers
 1..296
 /organism="Mus musculus"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone_image="2647683"
 /clone_lib="NCI CGAP Maml"
 /tissue_type="tumor, biopsy sample"
 /dev_stage="10 months, virgin"
 /lab_host="DH10B"
 /note="Organ: mammary; Vector: PCMV-SPORT6; site_1: SalI;
 site_2: NotI; Cloned unidirectionally. Primer: Oligo dr.
 library constructed by Life Technologies. Investigator
 providing samples: Gilbert Smith, NIH"
 BASE COUNT 62 a 71 c 95 g 68 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 0.00367 Length: 296
 Score: 14.00 Matches: 14
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.27% Indels: 0
 DB: 10 Gaps: 0
 US-09-397-967-16 (1-1099) x AM212845 (1-296)
 QY 1062 ProThrGluValGlnGluMetGlnLeuGlyTrpAlaPro 1075
 ||||||||||||||||||||||||||||||||||||||||||||
 Db 295 CCCACCGAGGTTCAGAGCTCATGCACTGTGTGGCGCCC 254
 RESULT 87
 LOCUS AA008555 318 bp mRNA linear EST 25-JUL-1996
 DEFINITION mg86b10.r1 Soares mouse embryo NMME13.5 14.5 Mus musculus cDNA

BASE COUNT 156 a 85 c 104 g 104 t
ORIGIN Bluescript KS(+) after bulk excision from Lambda FLC I."

Alignment Scores:

Pred. No.: 0.00601 Length: 449
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservatave: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.27% Indels: 0
DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x BF711016 (1-449)

Oy 1005 ValTSPePhgElyValValLeuTyrGluLeuPheThrTyr 1018
DB 359 CTTGTGAGCTTGTGAGCTGTCTATACGACTTTTCACATAC 400

RESULT 89

BO780982/c

LOCUS 636 bp mRNA linear EST 26-JUL-2002
DEFINITION UI-R-FFO-cpd-c-08-0-UI.s1 UI-R-FFO Rattus norvegicus cDNA clone

ACCESSION BO780982
VERSION BO780982.1 GI:21989454

KEYWORDS

EST.

SOURCE

ORGANISM Norway rat.
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE 1 (bases 1 to 636)

AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.

TITLE Normalization and subtraction: two approaches to facilitate gene discovery

JOURNAL Genome Res. 6 (9), 791-806 (1996)

MEDLINE

COMMENT

Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: msoares@blue.weeg.uiowa.edu
Tissue Procurement: Jeff Stevens
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: DISTRIBUTION: Researchers may obtain clones
from Research Genetics (www.resgen.com).
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES

source

1. 636
Location/Qualifiers
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/clone="UI-R-FFO-cpd-c-08-0-UI"
/clone_lib="UI-R-FFO"
/tissue_type="Mixed tissues"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/note="Vector: pRT73-Pac (Pharmacia) with a modified
polylinker. Site_1: EcoR I; Site_2: Not I; UI-R-FFO is a
subtracted cDNA library containing the following tissue(s)
): Normal cartilage and SR-JMS Tumor Line. The
subtraction was made according to Bonaldo, Lennon and
Soares, Genome Research, 6:791-806, 1996. The
oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tags for these libraries are: CTAATGACG,
CATCTTGTGA,
TAG_LIB-UI-R-FFO

TAG_TISSUE="cartilage"
TAG_SEQ="CTAATGACG"
BASE COUNT 139 a 156 c 200 g 138 t 3 others
ORIGIN

Alignment Scores:

Pred. No.: 0.00909 Length: 636
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservatave: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.27% Indels: 0
DB: 14 Gaps: 0

US-09-397-967-16 (1-1099) x BO780982 (1-636)

Oy 1045 LeuLeuGluLeuLeuAlaGluGlyArgLeuProPro 1058
DB 609 CTTGTGAGCTGTGAGGAGGCGCCCTCTCCACACCC 568

RESULT 90

BO781477/c

LOCUS 683 bp mRNA linear EST 26-JUL-2002
DEFINITION UI-R-FFO-cpg-e-19-0-UI.s1 UI-R-FFO Rattus norvegicus cDNA clone

ACCESSION BO781477
VERSION BO781477.1 GI:21989494

KEYWORDS

EST.

SOURCE

ORGANISM Norway rat.
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE 1 (bases 1 to 683)

AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.

TITLE Normalization and subtraction: two approaches to facilitate gene discovery

JOURNAL Genome Res. 6 (9), 791-806 (1996)

MEDLINE

COMMENT

Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: msoares@blue.weeg.uiowa.edu
Tissue Procurement: Jeff Stevens
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: DISTRIBUTION: Researchers may obtain clones
from Research Genetics (www.resgen.com).
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES

source

1. 683
Location/Qualifiers
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/clone="UI-R-FFO-cpg-e-19-0-UI"
/clone_lib="UI-R-FFO"
/tissue_type="Mixed tissues"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/note="Vector: pRT73-Pac (Pharmacia) with a modified
polylinker. Site_1: EcoR I; Site_2: Not I; UI-R-FFO is a
subtracted cDNA library containing the following tissue(s)
): Normal cartilage and SR-JMS Tumor Line. The
subtraction was made according to Bonaldo, Lennon and
Soares, Genome Research, 6:791-806, 1996. The
oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tags for these libraries are: CTAATGACG,
CATCTTGTGA.

TAG_LIB=UI-R-FEO
TAG_TISSUE=cartilage
TAG_SEQ=CTAATGACG
BASE COUNT 151 a 169 c 213 g 149 t 1 others

ALIGNMENT SCORES:

Pred. No.: 0.0099 Length: 683
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.27% Indels: 0
Gaps: 0

US-09-397-967-16 (1-1099) x B0781477 (1-683)

OY 1045 LeuLeuGluLeuAlaGluGlyArgArgLeuProProPro 1058
|||||

Db 609 CTTTGAGACTGCTGGCAGAGGGCCGCTCCACACACC 568

RESULT 91

LOCUS B1643881 381 bp mRNA linear EST 11-SEP-2001
DEFINITION NXPV_126.D07_F NXPV (Nsf Xylem Planings wood Vertical) Pinus taeda
ACCESSION B1643881
VERSION B1643881.1 GI:1553385
KEYWORDS EST.

SOURCE lobliolly pine.
ORGANISM Pinus taeda

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Coniferosida; Coniferales; Pinaceae; Pinus; Pinus.
1 (bases 1 to 381)
AUTHORS Sederoff, R.

TITLE Molecular Basis of Wood Formation in the Pine Megagenome
JOURNAL Unpublished (2000)
COMMENT Contact: Johnson, Arthur
North Carolina State University
Tel: 919 515 7800
Fax: 919 515 7801
Email: ajohnson@unity.ncsu.edu

FEATURES

source Location/Qualifiers

1..381
/organism="Pinus taeda"
/strain="Coastal plain lobliolly pine from North Carolina"
/db_xref="taxon:3352"
/clone="NXPV_126.D07"
/clone_1id="NXPV (Nsf Xylem Planings wood Vertical)"
/tissue_type="Xylem"
/cell_type="Planings (secondary)"
/dev_stage="Transitional"
/lab_host="X11-Blue"
/note="Vector: Bluescript SK; Site_1: Eco RI; Site_2: XhoI
The library is from early (spring) secondary wood, taken
from a ten year old tree in the transitional phase. The
tree is a kind gift of the Westvaco Corporation. Secondary
xylem was harvested from the tree by peeling back the bark
and primary xylem and then removing the underlying tissue
with a block plane. NOTE: The sequences contain a 'cna'
adapter' between the EcoRI site and the start of the EST.
The adapter sequence is 'AATTGGCAGCAG'."
BASE COUNT 118 a 75 c 81 g 92 t 15 others
ORIGIN

ALIGNMENT SCORES:

Pred. No.: 0.0471 Length: 381
Score: 13.00 Matches: 13
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.18% Indels: 0
Gaps: 0

US-09-397-967-16 (1-1099) x B1643881 (1-381)

OY 1003 SerAspValTrpSerPheGlyValValLeuTYrGluLeu 1015
|||||

Db 140 AGTGAGCTTTGGAGCTTGAGTACTGTGTATGACACTA 178

RESULT 92

LOCUS BG509232/2 441 bp mRNA linear EST 28-NOV-2001
DEFINITION BG509232
Bg509232
Sac87H10.Y1 Gm-c1073 glycine max cDNA clone GENOME SYSTEMS CLONE
ID: Gm-c1073-355 5' similar to FR:081064 081064 HYPOTHETICAL 51.9
KDA PROTEIN. //, mRNA sequence.
ACCESSION BG509232
VERSION BG509232.1 GI:13479889
KEYWORDS EST.

SOURCE soybean.
ORGANISM Glycine max

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
1 (bases 1 to 441)
AUTHORS Shoemaker, R., Keim, P., Vodkin, L., Erpelidg, J., Corvett, V., Khanna
A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,
Wyllie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers
Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk
R., Rittler, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann
R., Waterston, R. and Wilson, R.

TITLE Public Soybean EST Project
JOURNAL Unpublished (1999)
COMMENT Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu

FEATURES

source Location/Qualifiers

1..441
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-c1073-355"
/clone_1id="Gm-c1073"
/tissue_type="seedlings induced for symptoms of SDS
(Sudden Death Syndrome) disease"
/dev_stage="2-3 weeks old"
/lab_host="DH10B"
/note="Vector: Bluescript II SK+; Site_1: EcoRI; Site_2:
XhoI; The cDNA library was constructed from mRNA isolated
from 2-3 week old seedlings that were induced for symptoms
of SDS (Sudden Death Syndrome) disease by the
translocation of culture filtrate of Fusarium solani f.
sp. glycines (Plant Cell Report 18:375-380). Cultivar
Williams 82 is susceptible to the disease SDS. Plant
tissue (expanded leaves, folded leaves, and new shoots)
were collected at 1, 6, 24, and 48 hrs. after inoculation
and their mRNA pooled equally for cDNA construction. The
library was prepared using the Stratagene Bluescript II
SK(+) library construction kit. Complementary DNA was
synthesized from mRNA using a primer consisting of a
poly(dT) sequence with an XhoI restriction site. EcoRI
adapters were ligated to the blunt-ended cDNA fragments
followed by XhoI digestion. The cDNA insert is protected
from XhoI digestion via methylation during first strand
synthesis. The cDNA fragments were directionally cloned
into the EcoRI-XhoI restriction site of the plasmid

vector. The ligated cDNA fragments were transformed into E.coli Electromax DH10B host cells. Plants were inoculated by Shuxian Li (Glen Hartman lab, University of Illinois). Library was constructed by Reena Phillip and Steve Clough (Lila Vodkin lab, University of Illinois)."

BASE COUNT 156 a 90 c 64 g 131 t
ORIGIN

Alignment Scores:

Pred. No.:	0.0561	Length:	441
Score:	13.00	Matches:	13
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.18%	Indels:	0
DB:	12	Gaps:	0

US-09-397-967-16 (1-1099) x BG509232 (1-441)

Oy 1003 SeraspvairpserphedglyvalleuTYrGluLeu 1015

DB 133 ACTGATGTCGACGCTTGTGCTTCTTATGACCTT 95

RESULT 93

BM813523 444 bp mRNA linear EST 05-MAR-2002
LOCUS BM813523
DEFINITION EST591616 BNR Medicago truncatula/Meloidogyne incognita mixed EST library cDNA clone pBNIR-10G16, mRNA sequence.

ACCESSION BM813523.1 GI:19149537

VERSION

Medicago truncatula/Meloidogyne incognita mixed EST library.

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

1 (bases 1 to 444)
Bird,D., Kolat,H., Samac,D., Town,C.D., Van Aken,S., Uterback,T.,
Cheung,F., Tsai,J. and Fraser,C.M.
ESTs from roots of Medicago truncatula after infection with the
nematode Meloidogyne incognita
Unpublished (2002)
Contact: Bird, DM
Plant Nematode Genetics Group
North Carolina State University
Box 7616, Raleigh, NC 27695, USA
Tel: 919 515 6813
Fax: 919 515 9500
Email: david_bird@ncsu.edu
TIGR sequence name: MTC04J44TK More information is available at:
www.medicago.org
Seq primer: SKmod (CTA GAA CTA gty gAT CC).
Location/Qualifiers
1..444
/organism="Medicago truncatula/Meloidogyne incognita mixed
EST library"
/cultivar="A17"
/db_xref="taxon:188702"
/clone="pBNIR-10G16"
/clone.lib="BNIR"
/tissue_type="3 week old roots"
/dev_stage="3 days after infection with Meloidogyne
incognita second stage larvae"
/lab_host="XLOLR"
/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA
was directionally ligated into the Unizap XR vector from
Stratagene and packaged using Gigapack III Gold packaging
extracts. Plasmids containing cDNA inserts were excised
from the recombinant lambda-Zap phage using Ex-assist
helper phage and propagated in XLOLR cells."

BASE COUNT

145 a 81 c 106 g 112 t

ORIGIN

Alignment Scores: 0.0565 Length: 444

Score: 13.00 Matches: 13
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.18% Indels: 0
DB: 14 Gaps: 0

US-09-397-967-16 (1-1099) x BM813523 (1-444)

Oy 1003 SeraspvairpserphedglyvalleuTYrGluLeu 1015

DB 24 AGTGATGTCGACGCTTGTGCTTCTTATGACCTT 62

RESULT 94

BI424599 500 bp mRNA linear EST 29-NOV-2001
LOCUS BI424599
DEFINITION sah53e02.Y1 Gm-cl036 glycine max cDNA clone GENOME SYSTEMS CLONE
ID: Gm-cl036-5188-5' similar to TR:081064 081064 HYPOTHETICAL 51.9
KDA PROTEIN. ; mRNA sequence.

ACCESSION BI424599.1 GI:15200920

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 500)
Shoemaker,R., Kelm,P., Vodkin,L., Eprelding,J., Corvett,V., Khanna
A., Bolla,B., Maria,M., Hillier,L., Kueba,T., Martin,J., Beck,C.,
Wyllie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
R., Ritter,E., Kohn,S., Shiu,T., Jackson,Y., Cardenas,M., McCann
R., Waterston,R. and Willson,R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: Resgen, Invitrogen Corp. 2130
South Memorial Parkway Huntsville, AL 35801 For further information
call: (800)-533-4363 or contact via email: cut@resgen.com
High quality sequence stop: 414.
Location/Qualifiers
1..500
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl036-5188"
/clone.lib="Gm-cl036"
/tissue_type="somatic embryos cultured on MSD 20"
/lab_host="DH10B"
/note="Vector: pSPOR1; Site_1: NotI; Site_2: SalI; This
cDNA library was constructed from mRNA isolated from
somatic embryos (age ranging from 2 months to 9 months)
cultured on MSD 20. The library was prepared using the
life technologies psperscript cDNA library construction
kit. Complementary DNA was synthesized from mRNA using a
poly (dT) sequence with a NotI restriction site. SalI
linkers adapters were ligated to the blunt-ended cDNA
fragments followed by NotI digestion. The cDNA fragments
were directionally cloned into the NotI-SalI restriction
site of the pSPOR1 vector. The ligated cDNA fragments
were transformed into E.coli Electromax DH10B host cells.
This library was constructed in the laboratory of Dr. Lila
Vodkin by Anu Khanna at the University of Illinois at
Urbana-Champaign. e-mail: l-vodkin@uiuc.edu"

FEATURES

source

BASE COUNT

130 a 90 c 136 g 144 t

ORIGIN

Alignment Scores:

Pred. No.:	0.0651	Length:	500
Score:	13.00	Matches:	13
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.18%	Indels:	0
DB:	13	Gaps:	0

US-09-397-967-16 (1-1099) x B1424599 (1-500)

QY 1003 SeraspValTrrpSerPheGlyValValLeuTyrGluLeu 1015
 |||
 DB 348 AGTGATGTGTGGAGCTTGTGTAGTCTTATGAGCTT 386

RESULT 95
 AM231037 562 bp mRNA linear EST 10-DEC-1999
 LOCUS uc70a12.y1 NCI-CGAP_Mam1 Mus musculus cDNA IMAGE:2647870 5'
 DEFINITION similar to SW:JAK2_MOUSE Q62120 TYROSINE-PROTEIN KINASE JAK2 ;'
 mRNA sequence.

ACCESSION AM231037 GI:6560269
 VERSION AM231037
 KEYWORDS EST.
 SOURCE house mouse;
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 562)
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaps-remail.nih.gov
 Tissue Procurement: Gilbert Smith, Ph.D.
 CDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 www.bio.lnl.gov/bdrr/image/image.html

MG1:1028322
 Seq primer: -40RP from Gibco
 High quality sequence stop: 402.

FEATURES
 source
 Location/Qualifiers
 1..562
 /organism="Mus musculus"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone_image="IMAGE:2647870"
 /clone_1lb="NCI-CGAP_Mam1"
 /tissue_type="tumor, biopsy sample"
 /dev_stage="10 months, virgin"
 /lab_host="DH10B"
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
 Site_2: NotI; Cloned unidirectionally. Primer: Oligo dr.
 library constructed by Life Technologies. Investigator
 providing samples: Gilbert Smith, NIH"

BASE COUNT 181 a 114 c 147 g 120 t

ORIGIN

Alignment Scores:

Pred. No.:	0.0748	Length:	562
Score:	13.00	Matches:	13
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.18%	Indels:	0
DB:	10	Gaps:	0

US-09-397-967-16 (1-1099) x AM231037 (1-562)

QY 1003 SeraspValTrrpSerPheGlyValValLeuTyrGluLeu 1015

DB 524 TCAGATGTGTGGAGCTTGGAGTGTCTCATACGAGCTT 562
 |||
 RESULT 96
 BE156215/c
 LOCUS BE156215
 DEFINITION OVO-HT0367-310100-102-h08 HT0367 Homo sapiens cDNA, mRNA sequence.
 ACCESSION BE156215
 VERSION BE156215.1 GI:8618936
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 145)
 AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Brites,M.R.,
 Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
 Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bata,G.S., Simpson,D.H.,
 Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
 M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
 Simpson,A.J.
 Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 2020263
 COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the PAPSP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=st2-QVO-HT0367-310
 100-102-H08&t3=2000-01-31&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 17
 High quality sequence stop: 145.
 Location/Qualifiers
 1..145
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_1lb="HT0367"
 /dev_stage="Adult"
 /note="Organ: head,neck; Vector: puc18; Site_1: SmaI;
 Site_2: SmaI; A mini-library was made by cloning products
 derived from ORSTES PCR (U.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the puc 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."

BASE COUNT 35 a 40 c 42 g 28 t

ORIGIN

Alignment Scores:

Pred. No.:	0.143	Length:	145
Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.09%	Indels:	0
DB:	10	Gaps:	0

US-09-397-967-16 (1-1099) x BE156215 (1-145)

QY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
 |||
 DB 91 GTGCACCGCGACCTGGCTGCTCGCACATCCTGTC 56

RESULT 97
 AA320481
 LOCUS AA320481 214 bp mRNA linear EST 19-APR-1997

ACCESSION	IMAGE:160664.5' similar to gb:M57464 PROTO-ONCOGENE TYROSINE-PROTEIN KINASE RECEPTOR RET PRECURSOR (HUMAN); mRNA sequence.					
VERSION	H24996					
KEYWORDS	H24996.1	GI:893895				
SOURCE	human.					
ORGANISM	Homo sapiens					
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.					
AUTHORS	1 (bases 1 to 270) Hiller,L., Clark,N., Dubque,T., Elliston,R., Hawkins,M., Holman Riffin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston ,R., Williamson,A., Woldmann,P. and Wilson,R. The Washu-Memcon EST Project Unpublished (1995)					
TITLE	Contact: Wilson RK					
JOURNAL	Washington University School of Medicine					
COMMENT	4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel.: 314 286 1800 Fax: 314 286 1810 Email: est@watson.wustl.edu Insert Size: 809 High quality sequence stops: 166 Source: IMAGE Consortium, LNL This clone is available royalty-free through LNL ; contact the IMAGE Consortium (info@image.lnl.gov) for further information. Insert length: 809 Std Error: 0.00 Seq primer: M13RP1 High quality sequence stop: 166. Location/Qualifiers 1..270					
FEATURES	/organism="Homo sapiens"					
Source	/db_xref="GDB:574726"					
	/cdi_xref="taxon:9606"					
	/clone="IMAGE:160664"					
	/clone_lib="Soares breast 3MBHst"					
	/sex="Female"					
	/dev_stage="adult"					
	/lab_host="DH10B (ampicillin resistant)"					
	/note="Organ: breast; Vector: pT73D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5 TGTTCACATCTGAATGGAGCGCCGCCCTTTTCTTTTTTTTTTTT 3'] , double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not and Eco RI sites of a modified pT73 vector (Pharmacia). Library went through one round of normalization to a Cot 20. library constructed by Bento Soares and M.Fatima Bonaldo."					
BASE COUNT	65 a	52 c	74 g	69 t	10 others	
ORIGIN						
Alignment Scores:	Pred. No.:	0.298	Length:	270		
	Score:	12.00	Matches:	12		
	Percent Similarity:	100.00%	Conservative:	0		
	Best Local Similarity:	100.00%	Mismatches:	0		
	Query Match:	1.09%	Indels:	0		
	DB:	14	Gaps:	0		
US-09-397-967-16 (1-1099) x H24996 (1-270)						
OY	941	VaLHLSArGAsPLeuAlAaLaarGsAnIleUeVal	952	.		
Db	29	GTTCArTGCGGACTTGGCAGCACCAAArCTCGGrA	64			
LOCUS	B1051048	294 bp	mRNA	linear	EST 15-JUN-2001	
DEFINITION	CM2-GNO283-100101-683-g12 GNO283 Homo sapiens CDNA, mRNA sequence.					
ACCESSION	B1051048					

```

VERSION      GR051048.1  GI:14458578
KEYWORDS     EST.
SOURCE       human.
ORGANISM     Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1 (bases 1 to 294)
AUTHORS      Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
             Nagai,M.A., da Silva,M. Jr., Zago,M.A., Bordin,S., Costa,F.R.,
             Goldman,G.H., Carvalho,A.F.S., Matsukuma,A., Bala,G.S., Simpson,D.H.,
             Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
             ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
             Simpson,A.J.
             Shotgun sequencing of the human transcriptome with ORF expressed
             sequence tags
             Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
JOURNAL      20202663
MEDLINE
COMMENT      Contact: Simpson A.J.G.
             Laboratory of Cancer Genetics
             Ludwig Institute for Cancer Research
             Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
             Brazil
             Tel.: +55-11-2704922
             Fax: +55-11-2707001
             Email: asimpson@ludwig.org.br
             This sequence was derived from the FAPESP/LICR Human Cancer Genome
             Project. This entry can be seen in the following URL
             (http://www.ludwig.org.br/scripts/gethtml2.pl?l=CM2&l2=CM2-GN0283-100101-683-g12&t3=2001-01-10&t4=1)
             Seq primer: puc 18 forward
             High quality sequence start: 14
             High quality sequence stop: 294.
FEATURES     Location/Qualifiers
             1..294
             /organism="Homo sapiens"
             /db_xref="taxon:9606"
             /clone_lib="GN0283"
             /dev_stage="Adult"
             /note="Organ: Placenta-normal; Vector: puc18; Site_1: SmaI
             ; Site_2: SmaI; A mini-library was made by cloning
             products derived from ORESSES PCR (U.S. Letters Patent
             application No. 196,716 - Ludwig Institute for Cancer
             Research) profiles into the pUC 18 vector. Reverse
             transcription of tissue mRNA and cDNA amplification were
             performed under low stringency conditions."
BASE COUNT   58 a 81 c 90 g 65 t
ORIGIN
Alignment Scores:
Pred. No.: 0.33 Length: 294
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 13 Gaps: 0
US-09-397-967-16 (1-1099) x BI051048 (1-294)
Oy 941 ValHisArgAspLeuAlaAlaArgAsnLeuVal 952
| | | | | | | | | | | | | | | | | | | | | |
Db 270 GTTCACCGCTGACCTGGCTGCCGCAACATCTCTGTC 235

RESULT 100
BB605196 308 bp mRNA EST 05-DEC-2000
LOCUS BB605196 RIKEN full-length enriched, 0 day neonate lung Mus
DEFINITION musculus cDNA clone EO30009B22 5', mRNA sequence.
ACCESSION BB605196
VERSION BB605196.1 GI:11556598
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1	(bases 1 to 308)	Mammalia; Eutheria; Rodentia; Sciurognath; Muridae; Murinae; Mus		
	Alizawa, K., Akahira, S., Akimura, T., Arai, A., Arakawa, T., Carninci, P.,			
	Henagaki, T., Hayatsu, N., Hirotsu, T., Hirozane, T., Hodojima, Y.,			
	Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kawai, J., Kodama, Y., Konno			
	H., Kusakabe, M., Matsuyama, T., Miyazaki, A., Nakamura, M., Nishi, K.,			
	Nomura, K., Numanaki, R., Okazaki, Y., Okido, T., Owa, C., Sakai, C.,			
	Sakai, K., Sasaki, D., Sato, K., Shibata, K., Shibata, C.,			
	Shiraki, T., Sogabe, Y., Suzuki, H., Tagawa, A., Takahashi, F., Tanaka			
	T., Toyota, T., Watabiki, A., Yamamura, T., Yasunishi, A., Yoshida, K.,			
	Yoshiki, A., Yamatsu, M. and Hayashizaki, Y.			
	RIKEN Mouse ESTs (Alizawa, K. et al. 2000)			
	Unpublished (2000)			
	Contact: Yoshhide Hayashizaki			
	RIKEN Genomic			
	Center for Genome Exploration Research Group, RIKEN Genomic			

BASE COUNT	76 a	79 c	93 g	60 t
ORIGIN				
Alignment Scores:				
Pred. No.:	0.349			Length: 308
Score:	12.00			Matches: 12
Percent Similarity:	100.00%			Conservative: 0
Best Local Similarity:	100.00%			Mismatches: 0
Query Match:	1.09%			Indels: 0
DB:	10			Gaps: 0

US-09-397-967-16 (1-1099) x BB605196 (1-308)

QY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
|||||
DB 122 GTGCACCGGGACCTGGCGCCGCAACATCCTGTC 157

Search completed: April 28, 2003, 21:38:48
Job time : 3067 secs

THIS PAGE BLANK (USPTO)

PD 20-JUN-1996.
 XX
 PF 15-DEC-1995; 95WO-US16435.
 XX
 PR 15-DEC-1994; 94US-0357598.
 XX
 PA (UTJO) UNIV JOHNS HOPKINS SCHOOL MED.
 XX
 PI Clavin CI, Safford MG, Small D;
 XX
 DR WPI: 1996-300568/30.
 DR P-PSDB: AAR96037.
 XX
 PT Protein tyrosine kinase, JAK3, protein and nucleic acid - used in
 PT the gene therapy of cellular proliferative diseases, e.g. leukaemia,
 PT aplastic anaemia etc.
 XX
 PS Claim 4; Page 43-47; 97pp; English.
 XX
 CC A cDNA clone (AAT30862) codes for JAK3 (AAR96037), a new member of the
 CC JAK family of non-receptor protein tyrosine kinases, that probably
 CC plays a role in growth factor modulated differentiation,
 CC proliferation and survival of haematopoietic stem/progenitor cells.
 CC It was derived from CD34+ mRNA obtd. from normal human bone marrow
 CC by PCR amplification using primers (see also AAT30863-64) based
 CC on highly conserved motifs from protein tyrosine kinase catalytic
 CC domains. The JAK3 coding sequence can be incorporated into a
 CC vector and used for prodn. of recombinant JAK3. It can also be
 CC used in gene therapy protocols for leukaemia, myelodysplasia,
 CC polthemia vera, thrombocytosis and aplastic anaemia, or to
 CC stimulate haematopoietic cell proliferation. The gene was
 CC localised to chromosome 19, band p12-13.1.
 XX
 SQ Sequence 3807 BP; 716 A; 1219 C; 1099 G; 773 T; 0 other;
 SQ
 Alignment Scores:
 Pred. NO.: 2.37e-48 Length: 3807
 Score: 62.00 Matches: 96
 Percent Similarity: 97.96% Conservative: 0
 Best Local Similarity: 97.96% Mismatches: 1
 Query Match: 5.64% Indels: 2
 DB: 17 Gaps: 0
 US-09-397-967-16 (1-1099) x AAT30862 (1-3807)
 QY 938 ArgArgcysValHisArgaspLeuAlaalaArgasnileuValaGluserglualahis 957
 |||||||
 Db 2985 CGCGGTCGCGGACACCGGACCTGCGCCCGCAACATCCTCGTGAGAGAGGACGACAC 3044
 QY 958 VallysilealaaspheglyLeuAlaLysleuLeuPro-leuglyLysaspyrTYrrva 977
 |||||||
 Db 3045 GTCAAGATCGCTGACTTCGCGCTAGCTAGCTGCTGCGCGCTTGAC-AAAGACTACTAGCT 3103
 QY 977 lValaArgluproglyInserProIlephetpTyrrAlaProgluSerleuSerAspAs 997
 |||||||
 Db 3104 GGTCCGGAGCGACGACGACCCCATTTCTGTGATGCCGCCGAATCCCTCTGGAGAA 3163
 QY 997 nilePheSerArgInsleAspValTrrpserPheglyValValleuTYrrgluLeuPheTh 1017
 |||||||
 Db 3164 CATTTCTCTGCGCAGTCAGACGCTGTGAGGCTTGGGGTCTCTCTGTACGAGCTCTTAC 3223
 QY 1017 rTYrCysAspLysSerCysSerProSerAlaGlulPheLeuArgMetMetgly 1034
 |||||||
 Db 3224 CTACTGCGACAAAGCTGACGCCCTCGCGCGAGTCTCTGCGGAGTATGGGA 3275
 RESULT 3
 AAS10808
 1D AAS10808 standard; cDNA; 795 BP.
 XX
 AC AAS10808;
 XX
 DT 24-JUN-2001 (first entry)
 XX

DE Human Janus kinase 3 (JAK3) partial cDNA, 3-2 primer complement.
 XX
 KW Human; Janus kinase 3; JAK3; ss; JAK/STAT inhibitor; 3-2 primer;
 KW signal transducer and activator of transcription; osteoarthritis;
 KW degenerative joint disease; rheumatoid arthritis; leprosy; asthma;
 KW cancer; tumour; leukaemia.
 XX
 OS Homo sapiens.
 XX
 PN WO200152892-A2.
 XX
 PD 26-JUL-2001.
 XX
 PF 22-JAN-2001; 2001WO-US02033.
 XX
 PR 24-JAN-2000; 2000US-0177872.
 PR 28-NOV-2000; 2000US-0723490.
 XX
 PA (GENZ) GENZYME CORP.
 XX
 PI Vastios G;
 XX
 DR WPI: 2001-465338/50.
 XX
 PT Use of inhibitors of Janus kinase/signal transducers and activators of
 PT transcription for inhibiting onset and progression of degenerative
 PT joint diseases or disorders such as osteoarthritis, rheumatoid
 PT arthritis
 XX
 PS Example 4; Fig 3; 55pp; English.
 XX
 CC The sequence is a partial sequence, designated "complement of
 CC clone 3-2 primer", obtained from a chondrocyte cDNA encoding
 CC human Janus kinase 3 (JAK3). The invention relates to the use of
 CC JAK/STAT (Janus kinase/signal transducer and activator of
 CC transcription) inhibitors other than debromymentadistine (DBM) and
 CC hymenaldistine (H) for inhibiting the progression or the likelihood of
 CC developing diseases involving cartilage degradation, and for regulating
 CC the expression of pro-inflammatory agents or cytokines in a chondrocyte,
 CC and cartilage degradng enzymes in a cell. A JAK3/STAT inhibitor of the
 CC invention is useful for inhibiting progression or likelihood of
 CC developing osteoarthritis or rheumatoid arthritis. The inhibitor is also
 CC useful for treating other JAK/STAT-mediated diseases or disorders,
 CC including T cell-mediated disorders, mast cell-mediated disorders,
 CC type 2 (cytokine hypersensitivity) disorders, B cell lymphoma, and
 CC myeloid diseases. T cell-mediated disorders include human T cell
 CC leukaemia/lymphoma virus (HTLV)-1, Sdzory's syndrome, c-rabl
 CC transformation, natural killer-like T cell lymphomas (NK-like tumours)
 CC and graft-vs-host disease; cytokine hypersensitivity disorders include
 CC leishmaniasis, leprosy, allergy and viral infections; mast cell-mediated
 CC disorders include hay fever, asthma, hives and anaphylaxis; and
 CC leukaemias and lymphomas include acute lymphocytic and lymphoblastic
 CC leukaemias, B cell lymphomas and leukaemias of myeloid origin. DBM and H
 CC are useful as therapeutic agents in cancers in which JAK3 plays a role
 CC in the initiation or progression of tumorigenesis.
 XX
 SQ Sequence 795 BP; 156 A; 255 C; 226 G; 158 T; 0 other;
 SQ
 Alignment Scores:
 Pred. NO.: 1.55e-38 Length: 795
 Score: 51.00 Matches: 51
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 4.64% Indels: 0
 DB: 22 Gaps: 0
 US-09-397-967-16 (1-1099) x AAS10808 (1-795)
 QY 858 ProAspGlnArgaspPheglInarglulInleuLysAlaLeuHisSerAsp 877
 |||||||
 Db 455 CCAGACGACGAGGAGGACTTTCACGCGGAGATTCAGATCCCAAGACGTCGACAGTAT 514
 QY 878 PheIleVallystyrrArgglyValserTYrrglYProgluArgInsleuArgleuVal 897

```

Db      515  TTCAATGTCAGAGTATGCTGATGCGCGCGCGCGAGAGCTGGCTGGTC 574
Oy      898  MetcIuTYrleuProSerGlyCysLeuArgasp 908
Db      575  ATGAGTACTGCGCCAGCGGCTGCTGGCGGAC 607

RESULT 4
AA010809
ID      AA010809 standard: cDNA: 778 BP.
AC      AA010809;
DT      24-OCT-2001 (first entry)
XX      Human Janus kinase 3 (JAK3) partial cDNA sequence, 3-2 primer 1.
XX      Human; Janus kinase 3; JAK3; ss; JAK/STAT inhibitor; 3-2 primer 1;
XX      signal transducer and activator of transcription; osteoarthritis;
XX      degenerative joint disease; rheumatoid arthritis; leprosy; asthma;
XX      cancer; tumour; leukaemia.
OS      Homo sapiens.
PN      WO200152892-A2.
PD      26-JUL-2001.
PE      22-JAN-2001; 2001WO-US02033.
XX      24-JAN-2000; 2000US-0177872.
PR      28-NOV-2000; 2000US-0723490.
XX      (GENZ ) GENZYME CORP.
PA      Vastios G;
XX      WPI: 2001-465338/50.
DR      Use of inhibitors of Janus kinase/signal transducers and activators of
PT      transcription for inhibiting onset and progression of degenerative
PT      joint diseases or disorders such as osteoarthritis, rheumatoid
PT      arthritis
XX      Example 4; Fig 3; 55pp; English.
PS      The sequence is a partial sequence, designated "clone 3-2 primer 1",
XX      obtained from a chondrocyte cDNA encoding human Janus kinase
XX      3 (JAK3). The invention relates to the use of JAK/STAT (Janus
XX      kinase/signal transducer and activator of transcription)
XX      inhibitors other than debromohymenialdisine (DBH) and
XX      hymenialdisine (H) for inhibiting the progression or the likelihood of
XX      developing diseases involving cartilage degradation, and for regulating
XX      the expression of pro-inflammatory agents or cytokines in a chondrocyte,
XX      and cartilage degrading enzymes in a cell. A JAK3/STAT inhibitor of the
XX      invention is useful for inhibiting progression or likelihood of
XX      developing osteoarthritis or rheumatoid arthritis. The inhibitor is also
XX      useful for treating other JAK/STAT-mediated diseases or disorders,
XX      including T cell-mediated disorders, mast cell-mediated disorders,
XX      type 2 (cytokine hypersensitivity) disorders, B cell lymphoma, and
XX      myeloid diseases. T cell-mediated disorders include human T cell
XX      leukaemia/lymphoma virus (HTLV)-1, Sdzory's syndrome, c-abl
XX      transfection, natural killer-like T cell lymphomas (NK-like tumours)
XX      and graft-vs-host disease; cytokine hypersensitivity disorders include
XX      leishmaniasis, leprosy, allergy and viral infections; mast cell-mediated
XX      disorders include hay fever, asthma, hives and anaphylaxis; and
XX      leukaemias and lymphomas include acute lymphocytic and lymphoblastic
XX      leukaemias, B cell lymphomas and leukemias of myeloid origin. DBH and H
XX      are useful as therapeutic agents in cancers in which JAK3 plays a role
XX      in the initiation or progression of tumorigenesis.
SQ      Sequence 778 BP; 131 A; 264 C; 220 G; 163 T; 0 other;

```

```

Alignment Scores:
Pred. NO.: 4.33e-28 Length: 778
Score: 40.00 Matches: 74
Percent Similarity: 97.37% Conservative: 0
Best Local Similarity: 97.37% Mismatches: 1
Query Match: 3.64% Indels: 2
DB: 22 Gaps: 0

US-09-397-967-16 (1-1099) x AA010809 (1-778)
Oy      938  ArgArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGluSerGluAlaHis 957
Db      103  CGCCGCTGCGTGCACCGGACCTGGCCGCCCAACATCTCTGTGAGAGCGAGCGACAC 162
Oy      958  ValIysIleAlaAspPheGlyLeuAlaLysLeuPro-LeuGlyLysAspTyrTyrVa 977
Db      163  GTCAAGATCGCTGACTTGGCTTACCTAAGCTGCTGCCGCTTGAC-AAAGACTACTAGCT 221
Oy      977  lValArgGluProGlyGlnSerProIlePheTrpTyrAlaProGluSerLeuSerAspAs 997
Db      222  GGTCCGCGAGCGAGCGCCAGAGCCCATTTTCTGTATGCCCCGAAATCCCTCTCGACAA 281
Oy      997  nIlePheSerArgGlnSerAspValTrpSerPheGlyValValLeu 1012
Db      282  CATCTTCTCTCGCAGTCAAGACGTCTGGAGCTTCGGGGGTCTCTG 327

RESULT 5
AA087512
ID      AA087512 standard: cDNA to mRNA; 165 BP.
AC      AA087512;
DT      22-NOV-1995 (first entry)
XX      Tyrosine kinase peptide coding sequence.
XX      Tyrosine kinase peptide; UT-7 cell; primer: polymerase chain reaction;
XX      PCR; cell differentiation; antisense drug; ds.
OS      Homo sapiens.
PN      JP07059569-A.
XX      07-MAR-1995.
PD      25-AUG-1993; 93JP-0210404.
XX      25-AUG-1993; 93JP-0210404.
PR      (ASAH ) ASAH KASEI KOGYO KK.
XX      WPI: 1995-135894/18.
DR      P-PSDB: AAR71394.
PT      New DNA sequence encoding a tyrosine kinase peptide - useful for
PT      evaluation and control of cell differentiation.
XX      Claim 1; Page 5; 5pp; Japanese.
PS      This sequence encodes a tyrosine kinase peptide which is derived
XX      from UT-7 cells. This sequence was isolated using the primer
XX      sequences given in AA087510-11. The tyrosine kinase peptide may be
XX      used for the evaluation and control of cell differentiation and is
XX      also useful for the development of drugs and antisense drugs.
SQ      Sequence 165 BP; 36 A; 56 C; 39 G; 34 T; 0 other;

Alignment Scores:
Pred. NO.: 3.61e-20 Length: 165
Score: 31.00 Matches: 31
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.82% Indels: 0

```

DB: 16 Gaps: 0

US-09-397-967-16 (1-1099) x AAO87512 (1-165)

OY 973 LysAspTyrTyrValValArgGluProGlyGlnSerProIlePheTrpTyrAlaProGlu 992
 |||
 Db 73 AAGAGCTACTACGTGTCGGCAGCCAGCCAGACCCCATTTGTGATGCCGCCGAA 132
 |||

OY 993 SerLeuSerAspAsnIlePheSerArgGlnSer 1003
 |||
 Db 133 TCCCTCTCGGACACATCTTCTCTCGCAGTCA 165
 |||

RESULT 6

AAS10806 standard; cDNA: 811 BP.

AAS10806:
 AAS10806:
 24-OCT-2001 (first entry)

DE Human Janus kinase 3 (JAK3) partial cDNA sequence, 3-2 KS.

XX
 XX
 KM Human; Janus kinase 3; JAK3; ss: JAK/STAT inhibitor; 3-2 KS;
 KM signal transducer and activator of transcription; osteoarthritis;
 KM degenerative joint disease; rheumatoid arthritis; leprosy; asthma;
 KM cancer; tumour; leukaemia.

XX
 OS Homo sapiens.

XX
 PN WO200152892-A2.

XX
 PD 26-JUL-2001.

XX
 PF 22-JAN-2001; 2001WO-US02033.

XX
 PR 24-JAN-2000; 2000US-0177872.

PR 28-NOV-2000; 2000US-0723490.

PA (GEN2) GENZYME CORP.

PI Vastios G;

XX
 XX WPI; 2001-465338/50.

XX
 DR
 XX
 PT Use of inhibitors of Janus kinase/signal transducers and activators of
 PT transcription for inhibiting onset and progression of degenerative
 PT joint diseases or disorders such as osteoarthritis, rheumatoid
 PT arthritis

XX
 PS Example 4; Fig 3; 55pp; English.

XX
 CC The sequence is a partial sequence from a chondrocyte cDNA encoding
 CC human Janus kinase 3 (JAK3). The invention relates to the use of
 CC JAK/STAT (Janus kinase/signal transducer and activator of
 CC transcription) inhibitors other than debromohymenaldisine (DBH) and
 CC hymenaldisine (H) for inhibiting the progression or the likelihood of
 CC developing diseases involving cartilage degradation, and for regulating
 CC the expression of pro-inflammatory agents or cytokines in a chondrocyte,
 CC and cartilage degrading enzymes in a cell. A JAK3/STAT inhibitor of the
 CC invention is useful for inhibiting progression or likelihood of the
 CC developing osteoarthritis or rheumatoid arthritis. The inhibitor is also
 CC useful for treating other JAK/STAT-mediated diseases or disorders,
 CC including T cell-mediated disorders, mast cell-mediated disorders,
 CC type 2 (cytokine hypersensitivity) disorders, mast cell lymphoma, and
 CC myeloid diseases. T cell-mediated disorders include human T cell
 CC leukemia/lymphoma virus (HTLV)-1, Sdcory's syndrome, c-abl
 CC transformation, natural killer-like T cell lymphomas (NK-like tumours)
 CC and graft-vs-host disease; cytokine hypersensitivity disorders include
 CC leishmaniasis, leprosy, allergy and viral infections; mast cell-mediated
 CC disorders include hay fever, asthma, hives and anaphylaxis; and
 CC leukemias and lymphomas include acute lymphocytic and lymphoblastic
 CC leukemias, B cell lymphomas and leukemias of myeloid origin. DBH and H
 CC are useful as therapeutic agents in cancers in which JAK3 plays a role

CC in the initiation or progression of tumorigenesis.

XX
 XX Sequence 811 BP; 149 A; 254 C; 226 G; 182 T; 0 other;
 SQ

Alignment Scores:

Pred. No.:	Length:	811
Score:	29.00	Matches: 29
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	2.64%	Indels: 0
DB:	22	Gaps: 0

US-09-397-967-16 (1-1099) x AAS10806 (1-811)

OY 1006 TrpSerPheGlyValValLeuTyrGluLeuPheTrpTyrCysAspLysSerCysSerPro 1025
 |||
 Db 168 TGGACCTTCGGGGGTCCTCTTACGAGCTCTTACCTACTGCGACAAAGCTGACGCC 227
 |||

OY 1026 SerAlaGluPheLeuValArgMetMetGly 1034
 |||
 Db 228 TCGGCCGAGTTCCTGCGGATGATGGA 254
 |||

RESULT 7

ABN21423 standard; cDNA: 347 BP.

ABN21423:
 ABN21423:
 24-JUN-2002 (first entry)

DE Human ORFX polynucleotide sequence SEQ ID NO:11323.

XX
 XX
 DE Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;
 KM hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;
 KM degenerative disorder; osteoarthritis; neurodegenerative disorder;
 KM cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
 KM hypertension; hypothyroidism; cholesterol ester storage disease;
 KM immune deficiency; immune disorder; infectious disease;
 KM autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
 KM myasthenia gravis; gene; ss.

XX
 OS Homo sapiens.

XX
 PN WO200192523-A2.

XX
 PD 06-DEC-2001.

XX
 PF 29-MAY-2001; 2001WO-US10836.

XX
 PR 30-MAY-2000; 2000US-206132P.

PR 29-AUG-2000; 2000US-228716P.

XX
 PA (CURA-) CURAGEN CORP.

XX
 PI Shimkets RA, Leach MD;

XX
 XX WPI; 2002-106308/14.

XX
 DR P-PSDB; ABP05671.

XX
 PT Novel human polypeptides and polynucleotides useful for diagnosing,
 PT preventing and treating cardiovascular disease, neurodegenerative,
 PT hyperproliferative disorders and autoimmune disorders

XX
 PS Disclosure; SEQ ID 11323; 1037pp; English.

XX
 CC The present invention describes substantially purified human proteins
 CC (referred to as open reading frame, ORFX, where x is 1-11491 (see Table 1
 CC in the specification)). ABN15762 to ABN27252 encode the human ORFX
 CC proteins given in ABP00010 to ABP11500. ORFX proteins are useful for
 CC treating or preventing a pathology associated with an ORFX-associated
 CC disorder in humans, and in the manufacture of a medicament for treating a
 CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide
 CC sequences can be used in gene therapy. ORFX sequences can be used in the

CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
CC psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,
CC osteoarthritis, neurodegenerative disorders, disorders related to organ
CC transplantation, cardiovascular diseases, diabetes mellitus, systemic
CC lupus erythematosus, hypertension, hypochylodism, cholesterol ester
CC storage disease, various immune deficiencies and disorders, infectious
CC diseases, autoimmune thyroiditis, myasthenia gravis, rheumatoid
CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
CC disease and autoimmune inflammatory eye disease. ORF proteins are also
CC useful for treating burns, incisions, ulcers, for treating osteoporosis,
CC bone degenerative disorders, or periodontal disease, and for gut
CC protection or regeneration and treatment of lung or liver fibrosis,
CC reperfusion injury in various tissues and conditions resulting from
CC systemic cytokine damage.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SO Sequence 347 BP; 67 A; 117 C; 90 G; 69 T; 4 other;

Alignment Scores:
Pred. No.: 4,59e-16 Length: 347
Score: 27.00 Matches: 27
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2,46% Indels: 0
DB: Gaps: 24

US-09-397-967-16 (1-1099) x ABN21423 (1-347)

Oy 1043 CysArgLeuIleuGluLeuAlaGluGlyArgArgLeuProProProProProPro 1062
Db 19 TGCCTCCCTCTGAGCTCTGCGAGGAGCGACGCTCCACGCTCCACCTGCCCC 78

Oy 1063 ThrGluValGlnGluLeuMet 1069
Db 79 ACCGAGGTTCAGGACGCTCATG 99

RESULT 8
AAK78095
ID AAK78095 standard; DNA; 3222 BP.
XX
AC AAK78095;
XX
DT 07-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:32907.
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cystostatic; gene therapy; vaccine; metastasis; ds.
XX
OS Homo sapiens.
XX
PN MO200157182-A2.
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01354.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.

PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0233397.
PR 14-SEP-2000; 2000US-0233398.
PR 14-SEP-2000; 2000US-0233399.
PR 14-SEP-2000; 2000US-0233401.
PR 14-SEP-2000; 2000US-0233401.
PR 14-SEP-2000; 2000US-0233401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.

PR 01-NOV-2000; 2000US-024617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249246.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Rosen CA, Barash SC, Ruben SM;
 XX WPI; 2001-483426/52.
 XX
 PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating cancers and
 PT metastasis -
 XX
 XX
 PS Disclosure; SEQ ID NO 32907; 3071pp + Sequence Listing; English.

CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/hematopoietic-related diseases, especially
 CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/hematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
 CC represent sequences used in the exemplification of the present invention.
 XX
 SQ Sequence 3222 BP; 721 A; 883 C; 895 G; 723 T; 0 other;
 Alignment Scores:
 Pred. No.: 2,35e-11 Length: 3222
 Score: 23.00 Matches: 23
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 2.09% Indels: 0
 DB: 22 Gaps: 0
 US-09-397-967-16 (1-1099) x AAK78095 (1-3222)
 QY 325 GUAAlaGluPheProGlyLeuProGluAlaLeuSerPheValAlaLeuValAspGlyTyr 344
 Db 380 GAGCCGAGTTCCCGAGGCGTCTGTGCTTCGTGGCGCTCGTGGAGCGCTAC 439
 QY 345 PheArgLeu 347
 Db 440 TTCCGCGCTG 448
 RESULT 9
 ID AAKN57677 standard; DNA; 65 BP.
 XX AAKN57677;
 AC AAKN57677;
 XX
 DT 15-JUL-2002 (first entry)
 XX
 DE Mouse spliced transcript detection oligonucleotide SEQ ID NO:30425.
 XX
 KW Human; mouse; rat; splice transcript; detection; RNA transcript;
 KM splice variant; transcriptome; oligonucleotide library; ss.
 XX
 OS Mus musculus.
 XX
 PN W0200210449-A2.
 XX
 XX
 PD 07-FEB-2002.
 XX
 XX
 PF 20-JUL-2001; 2001MO-IB01903.
 XX
 XX
 PR 28-JUL-2000; 2000US-221607P.
 PR 02-MAY-2001; 2001US-287724P.
 XX
 XX
 PA (COMP-) COMPUGEN INC.
 PI Shoshan A, Wasserman A, Mintz E, Mintz L, Falgler S;
 XX
 XX
 DR WPI; 2002-257383/30.
 XX
 XX
 PT New oligonucleotide libraries comprising oligonucleotides which
 PT selectively hybridize to mRNAs transcribed from a transcription unit of
 PT a genome, useful for detecting tissue-, pathology-, and
 PT developmental-specific genes -
 XX
 XX
 PS Example 1; SEQ ID 30425; 47pp; English.
 XX
 XX The present invention describes oligonucleotide libraries for detecting
 CC messenger RNAs that populate a (sub-)transcriptome, where the
 CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
 CC transcription units that populate a genome. The library comprises
 CC several oligonucleotides, each capable of hybridizing selectively to a
 CC set of messenger RNAs transcribed from a given transcription unit of
 CC the genome, which encodes one or more messenger RNA splice variants.
 CC The oligonucleotide libraries are useful for detecting mRNAs from a
 CC biological sample, in expression profiling studies, in qualitatively or

CC quantitatively characterizing the corresponding transcriptome, and in
 CC detecting RNA transcripts and splice variants of human or animal
 CC transcriptomes. The libraries may also be used as specialised mini
 CC libraries to detect transcripts of a sub-transcriptome under a
 CC particular biological or pathological state, and so allowing the
 CC detection of tissue- and pathology-specific genes such as those genes
 CC only expressed in specific tissue under a specific pathological
 CC condition; to detect developmental specific genes; and to detect RNA
 CC transcripts and splice variants of a transcriptome of a patient suffering
 CC from a particular disorder. ABN27253 to ABN39389 represent
 CC oligonucleotide sequences from rats, humans and mice, which are used in
 CC the exemplification of the present invention.
 CC N.B. The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at [ftp://wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences).

SO Sequence 65 BP: 16 A; 18 C; 17 G; 14 T; 0 other;

Alignment Scores:

Pred. No.:	Length:	Score:
4,81e-11	65	21.00
Percent Similarity:	Matches:	21
100.00%	Conservative:	0
Best Local Similarity:	Mismatches:	0
Query Match:	Indels:	0
1.91%	Gaps:	0
24		

US-09-397-967-16 (1-1099) x ABN57677 (1-65)

OY 739 LeuProlaLeuysTrpThrgluLeuAlaGlyLeuIlethrcInCysMeIaTyasp 758

Db 3 CTGCCCCCTCTCAATGGACAGAACTGGCGACTTATCACACAGTGCATGGGATGAT 62

OY 759 Pro 759

Db 63 CCT 65

RESULT 10

AA510804

AA510804 standard: cDNA: 109 BP.

AC AA510804;

DT 24-OCT-2001 (first entry)

DE Human Janus kinase 3 (JAK3) RT-PCR fragment #1.

KW Human: Janus kinase 3; JAK3; RT-PCR; ss: JAK/STAT inhibitor;
 KW signal transducer and activator of transcription; osteoarthritis;
 KW degenerative joint disease; rheumatoid arthritis; leprosy; asthma;
 KW cancer; tumour; leukaemia.

OS Homo sapiens.

PN WO200152892-A2.

PD 26-JUL-2001.

PF 22-JAN-2001; 2001WO-US02033.

PR 24-JAN-2000; 2000US-0177872.

PR 28-NOV-2000; 2000US-0723490.

PA (GENZ) GENZYME CORP.

PI Vasilos G;

DR WPI; 2001-465338/50.

PT Use of inhibitors of Janus kinase/signal transducers and activators of
 PT transcription for inhibiting onset and progression of degenerative
 PT joint diseases or disorders such as osteoarthritis, rheumatoid
 PT arthritis
 XX

PS Example 2; Fig 1; 55pp; English.

XX The sequence is obtained from an RT (reverse transcriptase)-PCR fragment
 CC from human Janus kinase 3 (JAK3). The invention relates to the use of
 CC JAK/STAT (Janus kinase/signal transducer and activator of
 CC transcription) inhibitors other than dephosphorylating (DBP) and
 CC hynminaldine (H) for inhibiting the progression or the likelihood of
 CC the expression of pro-inflammatory agents or cytokines in a chondrocyte,
 CC and cartilage degrading enzymes in a cell. A JAK3/STAT inhibitor of the
 CC invention is useful for inhibiting progression or likelihood of
 CC developing osteoarthritis or rheumatoid arthritis. The inhibitor is also
 CC useful for treating other JAK/STAT-mediated diseases or disorders,
 CC including T cell-mediated disorders, mast cell-mediated disorders,
 CC type 2 (cytokine hypersensitivity) disorders, B cell lymphoma, and
 CC myeloid diseases. T cell-mediated disorders include human T cell
 CC leukaemia/lymphoma virus (HTLV)-1, Sezori's syndrome, c-abl
 CC transformation, natural killer-like T cell lymphomas (NK-like tumours)
 CC and graft-vs-host disease; cytokine hypersensitivity disorders include
 CC leishmaniasis, leprosy, allergy and viral infections; mast cell-mediated
 CC disorders include hay fever, asthma, hives and anaphylaxis; and
 CC leukaemias, B cell lymphomas include acute lymphocytic and lymphoblastic
 CC leukaemias, B cell lymphomas and leukaemias of myeloid origin. DBP and H
 CC are useful as therapeutic agents in cancers in which JAK3 plays a role
 CC in the initiation or progression of tumorigenesis.

SO Sequence 109 BP: 22 A; 36 C; 34 G; 15 T; 2 other;

Alignment Scores:

Pred. No.:	Length:	Score:
5.54e-08	109	18.00
Percent Similarity:	Matches:	18
100.00%	Conservative:	0
Best Local Similarity:	Mismatches:	0
Query Match:	Indels:	0
1.64%	Gaps:	0
22		

US-09-397-967-16 (1-1099) x AA510804 (1-109)

OY 945 LeuAlaAlaArgAsnIleLeuValGluSerGluAlaHisValIysIleAlaasp 962

Db 56 TTGGCCGCCGAAACATCTCTGGAGAGCGACGACATCAAGATCCCTGAC 109

RESULT 11

AA510807

AA510807 standard: cDNA: 726 BP.

AC AA510807;

DT 24-OCT-2001 (first entry)

DE Human Janus kinase 3 (JAK3) partial cDNA sequence, 3-2 SK.

KW Human: Janus kinase 3; JAK3; ss: JAK/STAT inhibitor; 3-2 SK;
 KW signal transducer and activator of transcription; osteoarthritis;
 KW degenerative joint disease; rheumatoid arthritis; leprosy; asthma;
 KW cancer; tumour; leukaemia.

OS Homo sapiens.

PN WO200152892-A2.

PD 26-JUL-2001.

PF 22-JAN-2001; 2001WO-US02033.

PR 24-JAN-2000; 2000US-0177872.

PR 28-NOV-2000; 2000US-0723490.

PA (GENZ) GENZYME CORP.

PI Vasilos G;

DR WPI; 2001-465338/50.

PR 02-MAR-2000; 200005-0186350

PR	15-MAR-2000	200005-0198074
PR	17-MAR-2000	200005-0198076
PR	18-APR-2000	200005-0198123
PR	19-MAY-2000	200005-0205515
PR	07-JUN-2000	200005-0209467
PR	28-JUN-2000	200005-0211486
PR	30-JUN-2000	200005-0215135
PR	07-JUL-2000	200005-0216647
PR	07-JUL-2000	200005-0216680
PR	11-JUL-2000	200005-0217486
PR	11-JUL-2000	200005-0218290
PR	26-JUL-2000	200005-0220964
PR	14-AUG-2000	200005-0224519
PR	14-AUG-2000	200005-0224519
PR	14-AUG-2000	200005-0225256
PR	14-AUG-2000	200005-0225567
PR	14-AUG-2000	200005-0225570
PR	14-AUG-2000	200005-0225747
PR	14-AUG-2000	200005-0225757
PR	14-AUG-2000	200005-0225759
PR	14-AUG-2000	200005-0225759
PR	14-AUG-2000	200005-0226261
PR	22-AUG-2000	200005-0226188
PR	22-AUG-2000	200005-0227182
PR	23-AUG-2000	200005-0227009
PR	30-AUG-2000	200005-0228924
PR	01-SEP-2000	200005-0229287
PR	01-SEP-2000	200005-0229343
PR	01-SEP-2000	200005-0229344
PR	01-SEP-2000	200005-0229345
PR	05-SEP-2000	200005-0229509
PR	05-SEP-2000	200005-0229513
PR	06-SEP-2000	200005-0230437
PR	06-SEP-2000	200005-0230438
PR	06-SEP-2000	200005-0231242
PR	08-SEP-2000	200005-0231243
PR	08-SEP-2000	200005-0231244
PR	08-SEP-2000	200005-0231413
PR	08-SEP-2000	200005-0231414
PR	08-SEP-2000	200005-0232080
PR	08-SEP-2000	200005-0232081
PR	12-SEP-2000	200005-0231966
PR	14-SEP-2000	200005-0233397
PR	14-SEP-2000	200005-0233398
PR	14-SEP-2000	200005-0233399
PR	14-SEP-2000	200005-0234400
PR	14-SEP-2000	200005-0234401
PR	14-SEP-2000	200005-0233063
PR	14-SEP-2000	200005-0233064
PR	14-SEP-2000	200005-0233065
PR	21-SEP-2000	200005-0234273
PR	21-SEP-2000	200005-0234274
PR	25-SEP-2000	200005-0234997
PR	25-SEP-2000	200005-0234998
PR	26-SEP-2000	200005-0235484
PR	27-SEP-2000	200005-0235834
PR	27-SEP-2000	200005-0235836
PR	29-SEP-2000	200005-0236527
PR	29-SEP-2000	200005-0236527
PR	29-SEP-2000	200005-0236567
PR	29-SEP-2000	200005-0236568
PR	29-SEP-2000	200005-0236569
PR	29-SEP-2000	200005-0236570
PR	02-OCT-2000	200005-0236602
PR	02-OCT-2000	200005-0237037
PR	02-OCT-2000	200005-0237038
PR	02-OCT-2000	200005-0237039
PR	02-OCT-2000	200005-0237040

PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 01-NOV-2000; 2000US-0241826.
 PR 08-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-024474.
 PR 08-NOV-2000; 2000US-024475.
 PR 08-NOV-2000; 2000US-024476.
 PR 08-NOV-2000; 2000US-024477.
 PR 08-NOV-2000; 2000US-024478.
 PR 08-NOV-2000; 2000US-024523.
 PR 08-NOV-2000; 2000US-024524.
 PR 08-NOV-2000; 2000US-024525.
 PR 08-NOV-2000; 2000US-024526.
 PR 08-NOV-2000; 2000US-024527.
 PR 08-NOV-2000; 2000US-024528.
 PR 08-NOV-2000; 2000US-024532.
 PR 08-NOV-2000; 2000US-024609.
 PR 08-NOV-2000; 2000US-024610.
 PR 08-NOV-2000; 2000US-024611.
 PR 17-NOV-2000; 2000US-024613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249246.
 PR 17-NOV-2000; 2000US-0249247.
 PR 17-NOV-2000; 2000US-0249248.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 08-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 11-DEC-2000; 2000US-0251990.
 PR 05-JAN-2001; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0255978.
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Rosen CA, Barash SC, Ruben SM;
 XX WPI: 2001-483426/52.
 XX
 XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating cancers and
 PT metastasis -
 XX
 XX Disclosure: SEQ ID NO 32909; 3071bp + Sequence Listing: English.
 CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)

CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytosolic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patient's own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting
 CC the nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/hematopoietic-related diseases, especially
 CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/hematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
 CC represent sequences used in the exemplification of the present invention.
 XX
 SQ Sequence 2000 BP; 345 A; 609 C; 640 G; 406 T; 0 other;
 Alignment Scores:
 Pred. No.: 8,49e-07 Length: 2000
 Score: 18.00 Matches: 18
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.64% Indels: 0
 DB: 22 Gaps: 0
 US-09-397-967-16 (1-1099) x AAK78097 (1-2000)
 QY 1 MetAlaProPserGlnGluThrProLeuIleProGlnArgSerCysSerLeu 18
 Db 25 ATGGCAGCTCCAGTAGAGAGAGCGCCCTGATGCTTCAGCGTTCAGCAGCTC 78
 RESULT 13
 AAV61801
 ID AAV61801 standard; cDNA: 3435 BP.
 XX
 XX AAV61801;
 AC 20-JAN-1999 (first entry)
 XX
 DT 20-JAN-1999 (first entry)
 XX
 DE JAK2 protein encoding cDNA.
 XX
 KW JAK2 protein; cytoplasmic domain; beta subunit; screening; asthma;
 KW Interleukin; granulocyte macrophage colony stimulating factor; GM-CSF;
 KW IL-3; IL-5; human; ss.
 XX
 OS Homo sapiens.
 XX
 FH
 FT Key Location/Qualifiers
 CDS 10..3429 /*tag= a
 FT /*product= "JAK2 protein"
 FT
 XX
 XX MO9843087-A1.
 XX
 XX 01-OCT-1998.
 XX
 XX 23-MAR-1998; 98MO-US05387.
 XX
 XX 24-MAR-1997; 97US-0041511.
 XX
 PA (PHAA) PHARMACIA & UPJOHN CO.
 XX
 XX Kayes PS, Roberds SL;
 PI
 XX
 XX WPI: 1998-532151/45.
 DR P-PSDB; AAM76425.
 DR
 XX
 XX Screening for compounds useful for preventing or treating asthma -
 PT by determining if compounds inhibit binding of the JAK2 protein to
 PT e.g. IL-3, IL-5 or GM-CSF
 XX

PS Disclosure: Fig 22A-C; 112pp; English.

CC This cDNA encodes the full-length JAK2 protein. The N-terminal fragment of JAK2 protein can be used in the method of the invention of screening for compounds useful for treating or preventing asthma. The method 294 comprises contacting a molecule comprising at least the N-terminal 294 amino acid residues of the JAK2 protein, with another molecule comprising at least 13 membrane-proximal cytoplasmic amino acids of Interleukin (IL)-3, IL-5 or granulocyte macrophage-colony stimulating factor (GM-CSF) proteins in the presence of the candidate compound, and determining whether the first and the second molecules form a complex. If the compound inhibits complex formation, it can be used to treat asthma.

SQ Sequence 3435 BP; 1157 A; 598 C; 733 G; 947 T; 0 other;

Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
Score:	0.000112	3435	16	0	0	0
Percent Similarity:	16.00					
Best Local Similarity:	100.00%					
Query Match:	1.46%					

DB: 19

US-09-397-967-16 (1-1099) x AAV61801 (1-3435)

QY 1003 SerAspValTrpSerPheGlyValValLeuTyrGluLeuPheThrTyr 1018
 |||||
 DB 3106 TCAGATGTGTGGAGCTTGGAGTGTCTGTATGAACCTTTCACATAC 3153

RESULT 14

ID AAQ25307 standard; DNA: 3473 BP.

XX AAQ25307;

AC 18-FEB-1999 (first entry)

DT JAK2 encoding DNA.

DE JAK2 encoding DNA.

XX Phosphorylation: JAK1, JAK2; protein tyrosine kinase; human;

KW catalytic domain; SH2 domain; growth factor receptor; PTK; murine; ss.

XX Mus musculus.

OS Mus musculus.

XX Key 1.2961 Location/Qualifiers

FT CDS /tag= a

FT /product= JAK2

XX MO9210519-A.

PN 25-JUN-1992.

PD 26-NOV-1991; 91WO-US08889.

PF 28-NOV-1990; 90AU-0003594.

PR (LUDWIG) INST CANCER RES.

XX PA Harpur A, Wilks AF, Ziemiecki A;

XX WPI: 1992-234591/28.

DR P-PSDB; AAR25141.

PT Novel protein tyrosine kinase mol. - comprises multiple catalytic domains but no SH2 domain and is for phosphorylation of proteins

XX Claim 10; Fig 8; 50pp; English.

PS This sequence encodes the murine protein tyrosine kinase JAK2 (from CC Janus kinase). Northern analysis of JAK2 expression in a mouse CC demonstrated the presence of two mRNA transcripts (4.8 and 4.4 kb) CC The levels of these transcripts alter with respect to one another in

CC different tissues. The kidney, spleen and lung appear to express CC predominantly the larger form, whereas ovary, placenta, skeletal muscle CC and all murine cell lines analysed express both forms at equal levels. CC The difference in sizes may be due to differential polyadenylation CC sites. Both JAK2 and JAK1 are examples of a new subfamily or class CC of protein tyrosine kinase. These can be used in the phosphorylation CC of proteins, incorporation of labels and in the design of analogues, CC antagonists and agonists of JAK's.

SQ Sequence 3473 BP; 1128 A; 677 C; 781 G; 887 T; 0 other;

Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
Score:	0.000113	3473	16	0	0	0
Percent Similarity:	16.00					
Best Local Similarity:	100.00%					
Query Match:	1.46%					

DB: 13

US-09-397-967-16 (1-1099) x AAQ25307 (1-3473)

QY 1003 SerAspValTrpSerPheGlyValValLeuTyrGluLeuPheThrTyr 1018
 |||||
 DB 2674 TCAGATGTGTGGAGCTTGGAGTGTCTGTATGAACCTTTCACATAC 2721

RESULT 15

ID AAD22680 standard; cDNA: 3495 BP.

XX AAD22680;

AC AAD22680;

XX 26-FEB-2002 (first entry)

DT Murine JAK2 tyrosine kinase cDNA.

DE Murine JAK2 tyrosine kinase cDNA.

XX Murine; Src homology 2-Bbeta; SH2-Bbeta; neuroprotective; gene therapy; cell differentiation; nerve regeneration; angiogenesis; embryogenesis; CC cytosolic; antisense therapy; drug screening; cellular expression; CC immunological disease; neurological disease; apoptosis; diabetes; cancer; arthritis; JAK2 tyrosine kinase; ss.

KW Mus sp.

OS Mus sp.

XX Key 1.2982 Location/Qualifiers

FT CDS /tag= a

FT /product= "Murine JAK2 tyrosine kinase protein"

XX US6312941-B1.

PN 06-NOV-2001.

PD 26-NOV-1997; 97US-0980080.

PF 26-NOV-1997; 97US-0980080.

PR (UNMI) UNIV MICHIGAN.

XX PA Carter-Su C, Rui L, Karow DS;

XX WPI: 2002-024907/03.

DR P-PSDB; AAE13587.

PT New isolated nucleic acid molecule encoding full length Src homology 2 active SH2-Bbeta protein -

XX Example 1; Fig 3A; 71pp; English.

PS The invention relates to human Src homology 2 (SH2)-Bbeta protein and its CC DNA molecule. SH2-Bbeta is useful in drug screening assays designed to CC identify drugs that interfere with the specific binding of JAK2 tyrosine CC kinase which is a member of Janus family of tyrosine kinase-interacting

CC signaling molecule. SH2-Bbeta protein is useful for modulating,
 CC preferably reducing cellular expression or intracellular concentration or
 CC availability of active SH2-Bbeta. SH2-Bbeta is useful for treating a
 CC disease associated with undesirable cell growth, differentiation, and
 CC growth-factor/cytokine, preferably interleukin, more preferably growth
 CC hormone, platelet derived growth factor, nerve growth factor, epidermal
 CC cell movement. SH2-Bbeta molecule has important implications in cancer
 CC metastasis, nerve regeneration, angiogenesis and embryogenesis and is
 CC useful for preventing apoptosis, and treating diabetes, cancer,
 CC arthritis, immunological diseases, neurological diseases, etc.
 CC The invention also relates to compositions and methods for identifying
 CC cytokine, hormone and growth factor signaling pathway agonists and
 CC antagonists. Human SH2-Bbeta DNA is useful in detecting the presence of
 CC SH2-Bbeta genes in gene transcripts, for detecting or amplifying DNA's
 CC with substantial sequence similarity with SH2-Bbeta homologues and
 CC structural analogues and for gene therapy applications. The present
 CC sequence is murine JAK2 tyrosine kinase cDNA related to the invention.
 CC
 XX
 SQ Sequence 3495 BP; 1121 A; 688 C; 794 G; 892 T; 0 other;

Alignment Scores:
 Pred. No.: 0.000114 Length: 3495
 Score: 16.00 Matches: 16
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.46% Indels: 0
 DB: 24 Gaps: 0

US-09-397-967-16 (1-1099) x AAD22680 (1-3495)

OY 1003 SeraspvaltrpserphcglyvalleuylrgluLeuPheThrTyr 1018
 DB 2695 TCAGATGCTGCGAGCTTGGAGTGTCTATACGAACCTTTCACATAC 2742

RESULT 16

AA085412
 ID AA085412 standard; cDNA; 3629 BP.

XX AC AA085412;
 XX DT 05-OCT-1995 (first entry)

XX DE Murine JAK2 kinase coding sequence and flanking regions.

XX KM JAK family; protein tyrosine kinase; cytokine receptor; mouse;

XX KW phosphorylation; signal transduction; activation; ss.

XX OS Mus musculus.

XX FH Key Location/Qualifiers

XX FT CDS 94..3483

XX FT conflict /tag= a

XX FT /product= Jak2_kinase

XX FT /tag= b

XX FT /note= "Published partial sequence of Jak2 cDNA

XX FT (Harpur et al., Oncogene 7:1347-1353(1992))

XX FT /tag= c

XX FT /note= "Published partial sequence of Jak2 cDNA

XX FT (Harpur et al., Oncogene 7:1347-1353(1992))

XX FT /tag= d

XX FT /note= "Published partial sequence of Jak2 cDNA

XX FT (Harpur et al., Oncogene 7:1347-1353(1992))

XX FT /tag= e

XX FT /note= "Published partial sequence of Jak2 cDNA

FT 1593(T), 1602(G), 1623(T), 1642(G), 1657(C),
 FT 1728(G), 1743(C), 1755(C), 1770(A), 1809(G),
 FT 1816(G), 1821(C), 1857(A), 1878(T), 1935(A),
 FT 1938(A), 1963(T), 1974(G), 2025(T), 2035(G),
 FT 2079(C), 2082(C), 2085(C), 2253(A), 2259(G),
 FT 2283(A), 2285(C), 2433(G), 3453(C), 3579(C),
 FT conflict
 FT 2226
 FT /tag= d
 FT /note= "location of a 7 amino acid insert detected
 FT in previous studies, but not in the present
 FT study"
 FT
 FT conflict
 FT 3595..3619
 FT /tag= e
 FT /note= "three extra nucleotides (all A's) were
 FT present in previous studies at positions
 FT 3595, 3598 and 3619 in the 3'-UTR"

PN W09503701-A.

PN 09-FEB-1995.

PD 29-JUL-1994; 94WO-US08676.

PF 29-JUL-1993; 93US-0097997.

PR (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.

PI Ihle JN, Quelle FW, Silvennoinen O, Witthuhn BA;

XX WPI: 1995-081950/11.

DR P-PSDB: AAR70830.

XX Inhibiting a cellular response to a cytokine by inhibiting Jak

XX kinase - to treat diseases caused by excessive response to

XX cytokine, e.g. erythrocytosis and other cellular proliferative

XX diseases

XX Claim 29; Fig 1; 167bp; English.

XX Inhibiting the activity of a jak kinase (pref. JAK1, JAK2, JAK3 or

XX TYK2) in a eukaryotic cell is claimed as a method of inhibiting the

XX biological response of that cell to a cytokine (not IL-3 or

XX erythropoietin). The present sequence (murine JAK2 kinase) encodes a

XX amino acid positions 758-776. Antibodies which selectively bind the

XX epitope are able to bind JAK2 without interfering with the activity

XX of the kinase. Such antibodies are claimed and are useful for

XX detecting and extracting JAK2. The 71 nucleotide differences noted

XX between the present sequence and the published partial sequence

XX result in 9 amino acid changes.

XX SQ Sequence 3629 BP; 1142 A; 733 C; 842 G; 912 T; 0 other;

XX Alignment Scores:

XX Pred. No.: 0.000118 Length: 3629

XX Score: 16.00 Matches: 16

XX Percent Similarity: 100.00% Conservative: 0

XX Best Local Similarity: 100.00% Mismatches: 0

XX Query Match: 1.46% Indels: 0

XX DB: 16 Gaps: 0

US-09-397-967-16 (1-1099) x AA085412 (1-3629)

OY 1003 SeraspvaltrpserphcglyvalleuylrgluLeuPheThrTyr 1018
 DB 3196 TCAGATGCTGCGAGCTTGGAGTGTCTATACGAACCTTTCACATAC 3243

RESULT 17

AA066244
 ID AA066244 standard; cDNA; 3629 BP.

XX AC AA066244;

DT	19-FEB-2001	(first entry)
DE	Jak2 polynucleotide sequence.	
XX		
XX	Jak3; kinase; cytokine; cellular response; inhibition; jak2; ss;	
KW	cell proliferation; erythrocytosis.	
XX		
XX	Unidentified.	
OS		
XX	US6136595-A.	
PN		
XX		
PD	24-OCT-2000.	
XX		
PF	18-JUN-1996;	96US-0665574.
XX		
PR	29-JUL-1994;	94US-0282012.
XX	29-JUL-1993;	93US-0097997.
PR	09-SEP-1993;	93US-0118968.
XX		
PA	(SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.	
XX		
PI	Silvennoinen O, Witthuhn BA, Ihle J;	
XX		
DR	WPI: 2000-686080/67.	
XX	P-PSDB: AAB35719.	
PT	New DNA encoding Jak3 kinase is useful as cytokine regulator for	
XX	treating cell proliferation	
PS	Example 1; Fig 1; 100pp; English.	
XX		
CC	This invention relates to DNA encoding a murine Jak3 protein. The amino	
CC	acid sequence of the Jak3 protein is given in AAB35715. The Jak family	
CC	of kinases are involved in the cellular response to the binding of	
CC	cytokines to their respective receptors. Jak3 kinase mediated activation	
CC	of some cytokines through their phosphorylation in response to	
CC	cytokine-receptor binding. Inhibiting the activity of jak3 kinase (at the	
CC	nucleic acid level with antisense sequences or ribozymes, or at the	
CC	protein level with antibodies, kinase inhibitors etc.) is used to treat,	
CC	or diagnose, diseases caused by excessive secretion of certain cytokines,	
CC	e.g. excessive cell proliferation such as erythrocytosis. Alternatively,	
CC	the Jak3 protein and polynucleotide can be used to treat conditions	
CC	associated with defective Jak3 activity. The DNA sequence can be used to	
CC	produce recombinant Jak3 and this used to raise antibodies useful as	
CC	specific inhibitors or to detect or isolate Jak3 without interfering	
CC	with its enzymatic activity. The present sequence represents cDNA	
CC	encoding a Jak2 protein used in the isolation and characterisation of the	
CC	Jak3 protein of the invention.	
SO	Sequence 3629 BP; 1142 A; 732 C; 843 G; 912 T; 0 other;	
Alignment Scores:		
Pred. No.:	0.000118	Length: 3629
Score:	16.00	Matches: 16
Percent Similarity:	100.00%	Conservative: 0
Best local Similarity:	100.00%	Mismatches: 0
Query Match:	1.46%	Indels: 0
DB:	21	Gaps: 0
US-09-397-967-16 (1-1099) x AAC66244 (1-3629)		
QY	1003 SerAspValTTPSerPheGlyValValLeuTYGILdeuPheThrTYR 1018	
Db	3196 TCAGATGTGTGAGAGCTTTGGAGTGTTCATACGAACTTTTCACATAC 3243	
RESULT 18		
ID	AAD03607	
XX	AAD03607 standard; cDNA; 3629 BP.	
XX	AAD03607;	
XX	19-JUN-2001 (first entry)	

DE	Murine (Janus kinase 2) Jak2 cDNA.	Location/Qualifiers	
XX		1..93	
XX		/*tag= a	
KM	Mouse; Janus kinase 2; Jak2; therapy; cytokine; tyrosine kinase;	94..3483	
KW	antiproliferative; cytostatic; cell proliferative disorder;	/*tag= b	
KW	cellular response; erythrocytosis; ss.	/*product= "Murine (Janus kinase 2) Jak2 protein"	
XX		523	
OS	Mus sp.	/*tag= c	
XX		/note= "5' end of published partial Jak2 cDNA	
FT		sequence (Harpur et al.)"	
FT		replace (551..553, CCC)	
FT		/*tag= d	
FT		replace (1089, T)	
FT		/*tag= e	
FT		replace (1103, C)	
FT		/*tag= f	
FT		replace (1111, G)	
FT		/*tag= g	
FT		replace (1119, G)	
FT		/*tag= h	
FT		replace (1122, C)	
FT		/*tag= i	
FT		replace (1128, C)	
FT		/*tag= j	
FT		replace (1131, G)	
FT		/*tag= k	
FT		replace (1134, G)	
FT		/*tag= l	
FT		replace (1137, C)	
FT		/*tag= m	
FT		replace (1140, G)	
FT		/*tag= n	
FT		replace (1143, G)	
FT		/*tag= o	
FT		replace (1146, C)	
FT		/*tag= p	
FT		replace (1188, T)	
FT		/*tag= q	
FT		replace (1194, G)	
FT		/*tag= r	
FT		replace (1230, G)	
FT		/*tag= s	
FT		replace (1245, T)	
FT		/*tag= t	
FT		replace (1260, T)	
FT		/*tag= u	
FT		replace (1266, C)	
FT		/*tag= v	
FT		replace (1272, T)	
FT		/*tag= w	
FT		replace (1275, Q)	
FT		/*tag= x	
FT		replace (1293, T)	
FT		/*tag= y	
FT		replace (1305, T)	
FT		/*tag= z	
FT		replace (1323, C)	
FT		/*tag= aa	
FT		replace (1341, A)	
FT		/*tag= ab	
FT		replace (1344, A)	
FT		/*tag= ac	
FT		replace (1359, G)	
FT		/*tag= ad	
FT		replace (1365, Q)	
FT			


```

FT      conflict
FT      /*tag- ae
FT      replace (1368, T)
FT      /*tag- af
FT      replace (1374, T)
FT      /*tag- ag
FT      replace (1401, C)
FT      /*tag- ah
FT      replace (1413, C)
FT      /*tag- ai
FT      replace (1431, T)
FT      /*tag- aj
FT      replace (1453, Q)
FT      /*tag- ak
FT      replace (1476, G)
FT      /*tag- al
FT      replace (1488, T)
FT      /*tag- am
FT      replace (1511..1512, GT)
FT      /*tag- an
FT      replace (1578, C)
FT      /*tag- ao
FT      replace (1590, T)
FT      /*tag- ap
FT      replace (1593, T)
FT      /*tag- aq
FT      replace (1602, G)
FT      /*tag- ar
FT      replace (1623, T)
FT      /*tag- as
FT      replace (1642, G)
FT      /*tag- at
FT      replace (1657, C)
FT      /*tag- au
FT      replace (1728, G)
FT      /*tag- av
FT      replace (1743, C)
FT      /*tag- aw
FT      replace (1755, C)
FT      /*tag- ax
FT      replace (1770, A)
FT      /*tag- ay
FT      replace (1809, G)
FT      /*tag- az
FT      replace (1816, G)
FT      /*tag- ba
FT      replace (1821, C)
FT      /*tag- bb
FT      replace (1857, A)
FT      /*tag- bc
FT      replace (1878, T)
FT      /*tag- bd
FT      replace (1935, A)
FT      /*tag- be
FT      replace (1938, A)
FT      /*tag- bf
FT      replace (1963, T)
FT      /*tag- bg
FT      replace (1974, G)
FT      /*tag- bh
FT      replace (2025, T)
FT      /*tag- bi
FT      replace (2055, G)
FT      /*tag- bj
FT      replace (2079, C)
FT      /*tag- bk
FT      replace (2082, C)
FT      /*tag- bl
FT      /*tag-
FT      misc_feature
FT      2226
FT      /*tag- bm
FT      /*note- "Nucleotides corresponding to 7 amino acid insert"
FT      replace (2253, A)
FT      /*tag- bn
FT      replace (2259, G)

```

```

FT      conflict
FT      /*tag- bo
FT      replace (2283, A)
FT      /*tag- bp
FT      replace (2285, C)
FT      /*tag- bq
FT      replace (2433, G)
FT      /*tag- br
FT      replace (3453, C)
FT      /*tag- bs
FT      3'UTR
FT      3484..3629
FT      /*tag- bt
FT      replace (3579, C)
FT      /*tag- bu
FT      replace (3595, A)
FT      /*tag- bv
FT      replace (3599, A)
FT      /*tag- bw
FT      replace (3620, A)
FT      /*tag- bx
FT      conflict
FT      US6210654-B1.
FT      03-APR-2001.
FT      08-OCT-1997; 97US-0946994.
FT      18-JUN-1996; 96US-0665574.
FT      29-JUL-1993; 93US-0097997.
FT      (SJD- ) ST JUDE CHILDREN'S HOSPITAL.
FT      Ihle J, Wittehu BA, Quelle FW, Silvenoinen O;
FT      WPI: 2001-265367/27.
FT      P-PSDB: AAE00352.
FT      Modulating a biological response mediated by Jak kinase 2 activation to
FT      a cytokine, useful for treating excessive proliferation of eukaryotic
FT      cells, comprises inhibiting or enhancing tyrosine kinase activity of
FT      Jak kinase in the cell.
FT      Example 1; Fig 1; 100pp: English.
FT      The present sequence is a cDNA encoding murine (Janus kinase 2) Jak2
FT      tyrosine kinase. Jak2 sequence has a 600 amino acid long N-terminus that
FT      lacks obvious SH2 (Src homology 2) and SH3 domains. Following this is a
FT      kinase related domain (domain 2) and a carboxyl kinase domain (domain
FT      1). Jak kinases mediate cytokine activity through their tyrosine
FT      Alignment Scores:
FT      Pred. No.: 0.000118 Length: 3629
FT      Score: 16.00 Matches: 16
FT      Percent Similarity: 100.00% Conservative: 0
FT      Best Local Similarity: 100.00% Mismatches: 0
FT      Query Match: 1.46% Indels: 0
FT      DB: 22 Gaps: 0
FT      US-09-397-967-16 (1-1099) x AAD03607 (1-3629)
FT      QY 1003 SeraspValITrpSerPheGlyValValLeuTyrgluLeupheThyr 1018
FT      Db 3196 TCAGATGTGTGAGAGCTTGGAGTGTCTCATACGAACCTTTCACATAC 3243
FT      RESULT 19
FT      AAX80971
FT      ID AAX80971 standard; CDNA: 4482 BP.
FT      AC AAX80971;
FT      XX
FT      XX 03-SEP-1999 (first entry)
FT      DE Human JAK2 kinase encoding cDNA.
FT      XX

```

KW JAK2 kinase; arteriosclerosis; asthma; bronchitis; emphysema; psoriasis;
 KW inflammatory bowel disease; inflammation; osteoarthritis; oncogenesis;
 KW rheumatoid arthritis; septic shock; systemic lupus erythematosus;
 KW leukemia; human; ss.
 OS Homo sapiens.
 XX
 XX US5914393-A.
 PN 22-JUN-1999.
 PD
 XX 05-DEC-1995; 95US-0567508.
 PF
 XX 05-DEC-1995; 95US-0567508.
 PR
 XX (INCY-) INCYTE PHARM INC.
 PA
 XX Coleman R, Stuart SG;
 PI
 XX WPI; 1999-384188/32.
 DR P-PSDB; AAY21698.
 XX
 PT Polypeptides and amino acids useful for modulating human jak2 kinase
 PT activity
 XX
 XX Claim 2; Columns 23-28; 37pp; English.
 PS
 CC This cDNA encodes a human JAK2 kinase polypeptide. Host cells transfected
 CC with recombinant jak2 kinase nucleic acid are used for the recombinant
 CC production of the protein. Purified JAK2 may be used to produce
 CC antibodies or identify antagonists or inhibitors of JAK2. JAK2, anti-JAK2
 CC antibodies and JAK2 antagonists or inhibitors may be used to treat,
 CC prevent or diagnose conditions associated with altered or uncontrolled
 CC JAK2 expression, e.g. arteriosclerosis, asthma, bronchitis, emphysema,
 CC inflammatory bowel disease, inflammation, leukemia, oncogenesis,
 CC osteoarthritis, psoriasis, rheumatoid arthritis, septic shock and
 CC systemic lupus erythematosus.
 CC
 XX
 SQ Sequence 4482 BP; 1477 A; 795 C; 953 G; 1249 T; 8 other;
 SO
 Alignment Scores:
 Pred. No.: 0.000144 Length: 4482
 Score: 16.00 Matches: 16
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.46% Indels: 0
 Gaps: 0
 DB: 20
 US-09-397-967-16 (1-1099) x AAX80971 (1-4482)
 OY 1003 SerAspValTrpSerPheGlyValValLeuTyrluLeuPheThrTYR 1018
 Db 3442 TCAGATGTTGGAGCTTGGAGTGTCTGTATGACTTTTCACATAC 3489
 RESULT 20
 AAZ58947
 ID AAZ58947 standard; cDNA; 4482 BP.
 XX
 AC AAZ58947;
 XX
 DT 03-MAY-2000 (first entry)
 DE Human JAK2 kinase (HJAK2) encoding cDNA.
 XX
 DE Human JAK2 kinase (HJAK2) encoding cDNA.
 XX
 KW Janus family nonreceptor protein-tyrosine kinase-2; JAK2; HJAK2; human;
 KW signal transduction; arteriosclerosis; asthma; bronchitis; emphysema;
 KW inflammatory bowel disease; leukemia; oncogenesis; osteoarthritis;
 KW psoriasis; rheumatoid arthritis; systemic lupus erythematosus;
 KW cytostatic; osteopathic; dermatological; antibacterial; septic shock;
 KW immunosuppressive; ss.
 XX
 OS Homo sapiens.
 XX

PN US6019966-A.
 XX
 XX 01-FEB-2000.
 PD
 XX 19-NOV-1998; 9805-0196480.
 PE
 XX 05-DEC-1995; 95US-0567508.
 PR
 XX (INCY-) INCYTE PHARM INC.
 PA
 XX Stuart SG, Coleman R;
 PI
 XX WPI; 2000-146859/13.
 DR P-PSDB; AAY77552.
 XX
 PT Human Janus family nonreceptor protein-tyrosine kinases useful as
 PT diagnostic reagents and for preventing, diagnosing and treating
 PT diseases such as arteriosclerosis, asthma and leukemia -
 XX
 XX Example 5; Fig 1A-F; 33pp; English.
 PS
 CC This cDNA encodes a human Janus family nonreceptor protein-tyrosine
 CC kinase-2 (JAK2) polypeptide (HJAK2). The JAK2 polypeptides may be used
 CC as diagnostic reagents as they react with a range of target proteins
 CC including growth hormone, prolactin, erythropoietin and cytokine
 CC receptors. They may also be used for the production of antibodies
 CC specific for JAK2, which may be used to inhibit its activity and prevent
 CC or treat disorders associated with over expression of JAK2. Conversely,
 CC the JAK2 polypeptide may be administered to supplement the patients own
 CC production and counter mutations that may lead to the expression of an
 CC inactive enzyme. The protein may also be used to screen candidate
 CC reagents for modulators of JAK2 function. The antagonists and antibodies
 CC bind to the JAK2 protein and prevent the transfer of high energy
 CC phosphate molecules, therefore blocking signal transduction. Disorders
 CC that may be treated by administration of JAK2 polypeptides, anti-JAK2
 CC antibodies and the agonists and/or antagonists, include arteriosclerosis,
 CC asthma, bronchitis, emphysema, inflammatory bowel disease, leukemia,
 CC oncogenesis, osteoarthritis, psoriasis, rheumatoid arthritis, septic
 CC shock and systemic lupus erythematosus.
 CC
 XX
 SQ Sequence 4482 BP; 1477 A; 795 C; 953 G; 1249 T; 8 other;
 SO
 Alignment Scores:
 Pred. No.: 0.000144 Length: 4482
 Score: 16.00 Matches: 16
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.46% Indels: 0
 Gaps: 0
 DB: 21
 US-09-397-967-16 (1-1099) x AAZ58947 (1-4482)
 OY 1003 SerAspValTrpSerPheGlyValValLeuTyrluLeuPheThrTYR 1018
 Db 3442 TCAGATGTTGGAGCTTGGAGTGTCTGTATGACTTTTCACATAC 3489
 RESULT 21
 AAD24311
 ID AAD24311 standard; DNA; 5117 BP.
 XX
 AC AAD24311;
 XX
 DT 07-MAR-2002 (first entry)
 DE Human JAK2 (Janus kinase) DNA.
 XX
 DE Human JAK2 (Janus kinase) DNA.
 XX
 KW Human; SOCS; suppressor of cytokine signalling; autoimmune disorder; JAK;
 KW Janus kinase; signal transducer and activator of transcription; Stat;
 KW rheumatoid arthritis; Wegener's granulomatosis; chronic active hepatitis;
 KW Crohn's disease; hemolytic anaemia; nephrotic syndrome; dermatological;
 KW diabetes mellitus; thyroiditis; inflammatory bowel disease; nephrotoxic;

KM	Immunosuppressive; antiinflammatory; antirheumatic; antiarthritic;
KW	hepatotropic; ds.
XX	
OS	Homo sapiens.
XX	
FM	Key
FM	Location/Qualifiers
FT	478..3876
FT	/*tag= a
FT	/product= "human jak2 protein"
XX	
PN	MO200179555-A2.
XX	
PD	25-OCT-2001.
XX	
PF	13-APR-2001; 2001MO-US12131.
XX	
PR	14-APR-2000; 2000OUS-0549654.
XX	
PA	(MILL-) MILLENNIUM PHARM INC.
PI	Hancock WM, Ozkaynak E;
XX	
DR	WPI: 2002-034368/04.
XX	
XX	P-PSDB: AAEL5179.
PT	Monitoring transplant acceptance or autoimmune disease, useful e.g. for
PT	assessing therapy, comprises measuring levels of Stat or their
PT	inhibitors
PS	Example; Fig 8; 218bp; English.
XX	
CC	The present invention relates to a method for monitoring acceptance of a
CC	transplant or an autoimmune disease in a mammal. The method comprising
CC	determining the amount of at least one of Stat4 (signal transducer and
CC	activator of transcription), Stat6, SOCS1 (suppressor of cytokine
CC	signalling) or SOCS3 mRNA or protein in a sample of the transplant, taken
CC	from the host or an affected tissue sample. Stats are activated by
CC	receptor-associated Janus kinases (Jaks) which include Jak1, Jak2, Tyk2,
CC	Jak3. The method is used to determine if autoimmune disorders (systemic
CC	lupus erythematosus, glomerulonephritis, rheumatoid arthritis, multiple
CC	sclerosis, myasthenia gravis, haemolytic anaemia, nephrotic syndrome,
CC	thyroiditis, diabetes mellitus, Crohn's disease, inflammatory bowel
CC	disease and vasculitis) are being treated successfully and may be used
CC	to adjust treatment regimes. The present sequence is human Jak2 DNA.
XX	
SO	Sequence 5117 BP; 1623 A; 927 C; 1103 G; 1464 T; 0 other;
Alignment Scores:	
Pred. No.:	0.000163
Score:	16.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	1.46%
DB:	24
	Gaps: 0
US-09-397-967-16 (1-1099) x AAD24311 (1-5117)	
Oy	1003 SeraspValITPSeRphnaglyValIvalLeuTyTcGlulnephenthrTyR 1018
DB	3580 TCAGATGTTTGGACCTTGGAGTGGTCTCTATGCACTTTTCACATAC 3627
RESULT 22	
AKK78096	
ID	AAK78096 standard; DNA: 1498 BP.
XX	
AC	AAK78096:
XX	
DT	07-NOV-2001 (first entry)
XX	
DE	Human immune/haematopoietic antigen genomic sequence SEQ ID NO:32908.
XX	

KM	Human: Immune: haematopoietic; Immune/haematopoietic antigen; cancer	
KW	cytostatic; gene therapy; vaccine; metastasis; ds.	
XX		
OS	Homo sapiens.	
XX		
PN	WO200157182-A2.	
XX		
PD	09-AUG-2001.	
XX		
PF	17-JAN-2001; 2001WO-US01354.	
XX		
PR	31-JAN-2000; 2000US-0179065.	
PR	04-FEB-2000; 2000US-0180628.	
PR	24-FEB-2000; 2000US-0184664.	
PR	02-MAR-2000; 2000US-0186350.	
PR	16-MAR-2000; 2000US-0189874.	
PR	17-MAR-2000; 2000US-0190076.	
PR	18-APR-2000; 2000US-0198123.	
PR	19-MAY-2000; 2000US-0205515.	
PR	07-JUN-2000; 2000US-0209467.	
PR	28-JUN-2000; 2000US-0214886.	
PR	30-JUN-2000; 2000US-0215135.	
PR	07-JUL-2000; 2000US-0216647.	
PR	07-JUL-2000; 2000US-0216880.	
PR	11-JUL-2000; 2000US-0217487.	
PR	11-JUL-2000; 2000US-0217496.	
PR	14-JUL-2000; 2000US-0218290.	
PR	26-JUL-2000; 2000US-0220963.	
PR	26-JUL-2000; 2000US-0220964.	
PR	14-AUG-2000; 2000US-0224518.	
PR	14-AUG-2000; 2000US-0224519.	
PR	14-AUG-2000; 2000US-0225213.	
PR	14-AUG-2000; 2000US-0225214.	
PR	14-AUG-2000; 2000US-0225267.	
PR	14-AUG-2000; 2000US-0225267.	
PR	14-AUG-2000; 2000US-0225268.	
PR	14-AUG-2000; 2000US-0225270.	
PR	14-AUG-2000; 2000US-0225447.	
PR	14-AUG-2000; 2000US-0225757.	
PR	14-AUG-2000; 2000US-0225758.	
PR	14-AUG-2000; 2000US-0225759.	
PR	18-AUG-2000; 2000US-0226279.	
PR	22-AUG-2000; 2000US-0226681.	
PR	22-AUG-2000; 2000US-0226868.	
PR	22-AUG-2000; 2000US-0227182.	
PR	23-AUG-2000; 2000US-0227009.	
PR	30-AUG-2000; 2000US-0228924.	
PR	01-SEP-2000; 2000US-0229287.	
PR	01-SEP-2000; 2000US-0229343.	
PR	01-SEP-2000; 2000US-0229344.	
PR	01-SEP-2000; 2000US-0229345.	
PR	05-SEP-2000; 2000US-0229509.	
PR	05-SEP-2000; 2000US-0229513.	
PR	06-SEP-2000; 2000US-0230437.	
PR	06-SEP-2000; 2000US-0230438.	
PR	08-SEP-2000; 2000US-0231242.	
PR	08-SEP-2000; 2000US-0231243.	
PR	08-SEP-2000; 2000US-0231244.	
PR	08-SEP-2000; 2000US-0231413.	
PR	08-SEP-2000; 2000US-0231414.	
PR	08-SEP-2000; 2000US-0232080.	
PR	08-SEP-2000; 2000US-0232081.	
PR	12-SEP-2000; 2000US-0231968.	
PR	14-SEP-2000; 2000US-0232387.	
PR	14-SEP-2000; 2000US-0232398.	
PR	14-SEP-2000; 2000US-0232399.	
PR	14-SEP-2000; 2000US-0232400.	
PR	14-SEP-2000; 2000US-0232401.	
PR	14-SEP-2000; 2000US-0233063.	
PR	14-SEP-2000; 2000US-0233064.	
PR	14-SEP-2000; 2000US-0233065.	
PR	21-SEP-2000; 2000US-0234223.	
PR	21-SEP-2000; 2000US-0234274.	

PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.

PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-483426/52.
XX
PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and
PT metastasis
XX
PS Disclosure; SEQ ID NO 32908; 3071bp + Sequence Listing; English.
XX
CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patient's own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I) by inserting
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/hematopoietic-related diseases, especially
CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/hematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
CC represent sequences used in the exemplification of the present invention.
XX
SQ Sequence 1498 BP; 359 A; 412 C; 407 G; 320 T; 0 other;
SQ
Alignment Scores:
Pred. No.: 0.0041 Length: 1498
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.27% Indels: 0
DB: Gaps: 0
US-09-397-967-16 (1-1099) x AAK78096 (1-1498)
QY 563 SerpheleugluAlaAlaSerleuMetserGlnValSerTyr 576
|||
Db 1206 TCATTCCTGGAAGCAGCGAGCTGTGATGAGCCAGGTGTGTAC 1247
RESULT 23
AAQ49750
ID AAQ49750 standard; DNA; 151 BP.
XX
AC AAQ49750;
XX
DT 10-MAR-1994 (first entry)
XX
DE pTK gene LpTK3 partial sequence.
XX
DE pTK; protein tyrosine kinase; catalytic domain; g-kit; NGF;
KW nerve growth factor receptor; megakaryocyte; lymphocyte;
KW amplification; primer; polymerase chain reaction; PCR; ss.
XX
OS Homo sapiens.
XX
PN WO9315201-A.
XX
PD 05-AUG-1993.
XX
PF 22-JAN-1993; 93WO-US00586.
XX
PR 22-JAN-1992; 92US-0826935.
XX
PA (NMEW-) NEW ENGLAND DEACONESS HOSPITAL.

```

XX  Avraham H, Cowley S, Groopman J, Scadden D;
PI  WPI; 1993-320330/40.
XX
XX  New protein tyrosine kinase genes and proteins encoded by genes -
PI  are of human megakaryocytic origin
XX
XX  Claim 2; Fig 3B; 60pp; English.
XX
CC  PTK genes were identified using two sets of degenerative
CC  oligonucleotide primers: a first set which amplifies all PTK DNA
CC  segments (AA049743-44), and a second set which amplifies highly
CC  conserved sequences present in the catalytic domain of the c-kit
CC  subgroup of PTKs (AA049745-46). The PTK genes identified are described
CC  in AA049747-57 and AAR41897-02.
CC  The LPTKs are expressed in lymphocytic cells, as well as
CC  megakaryocytic cells. The LPTK3 expression prod. exhibited
CC  significant sequence homology with known protein tyrosine kinases
CC  of the NGF receptor family. The protein sequence corresp. to AA049750
CC  is claimed (claim 7) and stated as given in the specification, however
CC  is missing from the publication.
XX
SQ  Sequence 151 BP; 37 A; 44 C; 41 G; 29 T; 0 other;

Alignment Scores:
Pred. No.:      0.0379      Length:      151
Score:          12.00       Matches:      12
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match:    1.09%       Indels:      0
DB:             14         Gaps:        0

US-09-397-967-16 (1-1099) x AA049750 (1-151)

OY  941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
ID  |||||||
DB  1 GTGCACAGGAGATCTCGGGCTCGAACAATCTCTGTC 36

RESULT 24
AA03093
XX  AAT03093 standard; DNA; 151 BP.
XX
AC  AAT03093;
XX
DT  14-FEB-1996 (first entry)
XX
DE  Protein tyrosine-kinase LPTK3 DNA fragment.
XX
KW  Protein tyrosine-kinase; PTK; LPTK3; agonist; cell growth;
KW  differentiation; ss.
XX
OS  Homo sapiens.
XX
FH  Key      Location/Qualifiers
FT  CDS      1..150
FT           /*tag= a
FT
XX  MO9527061-A1.
XX
XX  12-OCT-1995.
XX
XX  04-APR-1995; 95WO-US04228.
XX
XX  04-APR-1994; 94US-0222616.
XX
XX  (GENE ) GENENTECH INC.
XX
PI  Bennett BD, Goeddel D, Lee JM, Matthews W, Tsai SP;
PI  Wood WI;
XX
XX  WPI; 1995-366160/47.
XX
XX  P-PSDB; AAR85932.

```

```

XX  Agonist antibodies which activate specific protein tyrosine
PI  kinase(s) - also activate chimeric proteins of kinase extracellular
PI  domain and Ig constant domain, useful for studying, and therapeutic
PI  modulation of, cell growth and differentiation
XX
XX  Disclosure; Page 37; 125pp; English.
XX
CC  DNA probes based on protein tyrosine-kinase (PTK) sequences were used
CC  to screen cDNA libraries to identify novel PTK genes. A LPTK3 gene
CC  fragment (AAT03093) was isolated from lymphocytic and megakaryocytic
CC  cell line libraries and encoded a peptide (AAR85932) showing homology
CC  to known PTKs. The gene fragment can be used to identify other
CC  new PTK genes, or to design drugs, peptides or antisense constructs
CC  that modulate PTK activity.
XX
SQ  Sequence 151 BP; 37 A; 44 C; 41 G; 29 T; 0 other;

Alignment Scores:
Pred. No.:      0.0379      Length:      151
Score:          12.00       Matches:      12
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match:    1.09%       Indels:      0
DB:             16         Gaps:        0

US-09-397-967-16 (1-1099) x AAT03093 (1-151)

OY  941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
ID  |||||||
DB  1 GTGCACAGGAGATCTCGGGCTCGAACAATCTCTGTC 36

RESULT 25
AAS10805
XX  AAS10805 standard; cDNA; 214 BP.
XX
AC  AAS10805;
XX
DT  24-OCT-2001 (first entry)
XX
DE  Human Janus kinase 3 (JAK3) RT-PCR fragment #2.
XX
KW  Human; Janus kinase 3; JAK3; RT-PCR; ss; JAK/STAT inhibitor;
KW  signal transducer and activator of transcription; osteoarthritis;
KW  degenerative joint disease; rheumatoid arthritis; leprosy; asthma;
KW  cancer; tumour; leukaemia.
XX
OS  Homo sapiens.
XX
XX  WO200152892-A2.
XX
XX  26-JUL-2001.
XX
XX  22-JAN-2001; 2001WO-US02033.
XX
XX  24-JAN-2000; 2000US-0177872.
XX
XX  28-NOV-2000; 2000US-0723490.
XX
XX  (GENE ) GENZYME CORP.
XX
XX  Vasios G;
XX
XX  WPI; 2001-465338/50.
XX
XX  Use of inhibitors of Janus kinase/signal transducers and activators of
XX  transcription for inhibiting onset and progression of degenerative
XX  joint diseases or disorders such as osteoarthritis, rheumatoid
XX  arthritis -
XX
XX  Example 3; Fig 2; 55pp; English.
XX
XX  The sequence is obtained from an RT (reverse transcriptase)-PCR fragment
XX  from human Janus kinase 3 (JAK3). The invention relates to the use of

```

JAK/STAT Janus kinase/signal transducer and activator of transcription) inhibitors other than debrisoquine/aldazine (DBH) and hydralazine (H) for inhibiting the progression or the likelihood of developing diseases involving cartilage degradation, and for regulating the expression of pro-inflammatory agents or cytokines in a chondrocyte, and cartilage degrading enzymes in a cell. A JAK3/STAT inhibitor of the invention is useful for inhibiting progression or likelihood of developing osteoarthritis or rheumatoid arthritis. The inhibitor is also useful for treating other JAK/STAT-mediated diseases or disorders, including T cell-mediated disorders, mast cell-mediated disorders, type 2 (cytokine hypersensitivity) disorders, B cell lymphoma, and myeloid diseases. T cell-mediated disorders include human T cell leukaemia/lymphoma virus (HTLV)-1, Sclery's syndrome, c-abl transformation, natural killer-like T cell lymphomas (NK-like tumours) and graft-vs-host disease; cytokine hypersensitivity disorders include leishmaniasis, leprosy, allergy and viral infections; mast cell-mediated disorders include hay fever, asthma, hives and anapylaxis; and leukaemias and lymphomas include acute lymphocytic and lymphoblastic leukaemias, B cell lymphomas and leukaemias of myeloid origin. DBH and H are useful as therapeutic agents in cancers in which JAK3 plays a role in the initiation or progression of tumourigenesis.

50 Sequence 214 BP; 34 A; 74 C; 62 G; 36 T; 8' other;

Alignment Scores:	
Pred. No.:	0.0526
Score:	12.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	1.09%
DB:	22
Length:	214
Matches:	12
Conservative:	0
Mismatches:	0
Indels:	0
Gaps:	0

US-09-397-967-16 (1-1099) x AAS10805 (1-214)

Db 128 CGCCGCTGCGTGACCGCGACTGGCCGCCGGAAC 163

RESULT 26

ID ABA50838 standard; DNA; 294 BP.

AC ABA50838;

DT 01-FEB-2002 (first entry)

DE Human breast cell single exon nucleic acid probe #95333.

Human; microarray; single exon probe; gene expression; breast; disease; cancer; ss.

OS Homo sapiens.

PN W0200157271-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US006662.

PR	04-FEB-2000;	2000US-0180312
PR	26-MAY-2000;	2000US-0207456
PR	30-JUN-2000;	2000US-0608408
PR	03-AUG-2000;	2000US-0632366
PR	21-SEP-2000;	2000US-0234687
PR	27-SEP-2000;	2000US-0236359
PR	04-OCT-2000;	2000GB-0024263

PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-496933/54.

XX

PT New spatially-addressable set of single exon nucleic acid probes,
 PT useful for measuring gene expression in sample derived from human
 PT breast, comprises number of single exon nucleic acid probes -
 XX
 PS Claim 4; SEQ ID NO 9533; 327PP + sequence listing; English.

PS Claim 4; SEQ ID NO 9533; 327pp + sequence listing; English.

CC The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human breast and BT 474 cells. The method involves contacting
CC the probes with a collection of detectably labelled nucleic acids
CC derived from mRNA of human breast, and then measuring the label
CC bound to each probe of the microarray. The probes are useful for
CC verifying the expression of regions of genomic DNA predicted to
CC encode proteins. They are useful for gene discovery, and for
CC determining predisposition and/or prognosing breast disease. Gene
CC expression analysis is useful for assessing the toxicity of chemical
CC agents on cells. The microarray of this invention presents a far greater
CC diversity of probes for measuring gene expression, with far less bias
CC than expressed sequence tag microarrays. The method is suitable for
CC rapid production of functional information from genomic sequence. The
CC present sequence is a single exon nucleic acid probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.

Sequence 294 BP; 69 A; 60 C; 79 G; 86 T; 0 other;

Alignment Scores:	
Pred. No.:	0.0708
Score:	12.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	1.09%
DB:	22
	Gaps: 0
	Indels: 0
	Mismatches: 0
	Conservative: 0
	Matches: 12
	length: 294

US-09-397-967-16 (1-1099) x ABA50838 (1-294)

Db 91 GTGCATCGTGATCTGGCCGCACGGAACATCCTGGTG 126

RESULT 27

ID	ABA68808	standard; DNA; 294 BP

- AC ABA68808;

DT 01-FEB-2002 (first entry)

DE Human foetal liver single exon nucleic acid probe #17113.

KW Human; foetal liver; gene expression; single exon nucleic acid probe; ss

05 Homo sapiens

PN W0200157277-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00669.

PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207455.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024265.

PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX

DR WPI: 2001-483447/52.

XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human fetal liver -
XX
PS Claim 4: SEQ ID NO 17113: 639pp + sequence listing; English.

CC The invention relates to a single exon nucleic acid probe for
CC measuring human gene expression in a sample derived from human foetal
CC liver. The single exon nucleic acid probes may be used for predicting,
CC measuring and displaying gene expression in samples derived from human
CC fetal liver. The present sequence is a single exon nucleic acid
CC probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 294 BP; 69 A; 60 C; 79 G; 86 T; 0 other;

Alignment Scores:
Pred. No.: 0.0708 Length: 294
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 22 Gaps: 0

US-09-397-967-16 (1-1099) x ABA68808 (1-294)

QY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
DB 91 GTGCATCGTGAATCGCCGACGACATCTCTGCTG 126

RESULT 28
ABA35766
ID ABA35766 standard; DNA: 294 BP.

AC ABA35766:
XX
DT 23-JAN-2002 (first entry)
XX

DE Probe #14232 for gene expression analysis in human heart cell sample.

XX
XX Human: gene expression; heart; microarray; vascular system; probe;
KW cardiovascular disease; hypertension; cardiac arrhythmia;
KW congenital heart disease; ss.
XX
OS Homo sapiens.
XX
PN WO200157274-A2.
XX
PD 09-AUG-2001.
XX

PF 30-JAN-2001; 2001WO-US00666.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI: 2001-488899/53.
XX

PT Single exon nucleic acid probes for analyzing gene expression in human
PT hearts -
XX
PS Claim 4: SEQ ID NO 14232; 530pp; English.

XX

CC The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart. The
CC present sequence is one such probe. The probes may be used for
CC predicting, measuring and displaying gene expression in samples derived
CC from the human heart via microarrays. By measuring gene expression, the
CC probes are useful for predicting, diagnosing, grading, staging, the
CC monitoring and prognosing diseases of the human heart and vascular system
CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
CC congenital heart disease.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 294 BP; 69 A; 60 C; 79 G; 86 T; 0 other;

Alignment Scores:
Pred. No.: 0.0708 Length: 294
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 22 Gaps: 0

US-09-397-967-16 (1-1099) x ABA35766 (1-294)

QY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
DB 91 GTGCATCGTGAATCGCCGACGACATCTCTGCTG 126

RESULT 29
AAK17149
ID AAK17149 standard; DNA: 294 BP.

AC AAK17149:
XX
DT 05-NOV-2001 (first entry)
XX

DE Human brain expressed single exon probe SEQ ID NO: 17140.

XX
XX Human: brain expressed exon; gene expression analysis; probe;
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KW epilepsy; cancer; ss.
XX
OS Homo sapiens.
XX
PN WO200157275-A2.
XX
PD 09-AUG-2001.
XX

PF 30-JAN-2001; 2001WO-US00667.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI: 2001-483446/52.
XX

PT Single exon nucleic acid probes for analyzing gene expression in human
PT brains -
XX
PS Example 4: SEQ ID NO: 17140; 650pp + Sequence Listing; English.

CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human

CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is one of the probes of the
CC invention.

Sequence 294 BP; 69 A; 60 C; 79 G; 86 T; 0 other;

Alignment Scores:

Alignment Scores:	0.0708	Length:	294
Pred. No.:	12.00	Matches:	12
Score:	100.00%	Conservative:	0
Percent Similarity:	100.00%	Mismatches:	0
Best Local Similarity:	100.00%	Indels:	0
Query Match:	1.09%	Gaps:	0
DB:	22		

US-09-397-967-16 (1-1099) x AAK17149 (1-294)

OY 941 VALHISARGSPLEUAlAlARqAsnIleUeVal 952

DB 91 GTGCATCGTGAATCTGCGCCGACGGAACATCTCTG 126

RESULT 30

AAK42934

AC AAK42934:

XX 06-NOV-2001 (first entry)

DE Human bone marrow expressed single exon probe SEQ ID NO: 17491.

XX Human: bone marrow expressed exon; gene expression analysis; probe;

KW microarray; cancer; Leukemia; lymphoma; myeloma; ss.

OS Homo sapiens.

PN WO200157276-A2.

XX 09-AUG-2001.

PD 30-JAN-2001; 2001WO-US00668.

XX 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

PI WPI; 2001-488901/53.

DR Human genome-derived single exon nucleic acid probes useful for

PT analyzing gene expression in human bone marrow

XX Example 4; SEQ ID NO: 17491; 658bp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid

CC probes which are derived from genomic sequences expressed in the human

CC bone marrow. They can be used to measure gene expression in bone marrow

CC samples, which may enable the improved diagnosis and treatment of cancers

CC such as lymphoma, leukemia and myeloma. The present sequence is one of

XX the probes of the invention.

XX Sequence 294 BP; 69 A; 60 C; 79 G; 86 T; 0 other;

Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.09%	Indels:	0
DB:	22	Gaps:	0

US-09-397-967-16 (1-1099) x AAK42934 (1-294)

OY 941 VALHISARGSPLEUAlAlARqAsnIleUeVal 952

DB 91 GTGCATCGTGAATCTGCGCCGACGGAACATCTCTG 126

RESULT 31

AAI23695

ID AAI23695 standard; DNA; 294 BP.

AC AAI23695;

XX 12-OCT-2001 (first entry)

DE Probe #13628 for gene expression analysis in human cervical cell sample.

XX Probe; human; microarray; gene expression; cervical epithelial cell;

KW cervical cancer; ss.

OS Homo sapiens.

PN WO200157278-A2.

XX 09-AUG-2001.

PD 30-JAN-2001; 2001WO-US00670.

XX 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

PI WPI; 2001-488901/53.

DR Human genome-derived single exon nucleic acid probes useful for

PT analyzing gene expression in human cervical epithelial cells

XX Claim 25; SEQ ID No 13628; 487bp; English.

XX The present invention relates to human single exon nucleic acid probes

CC (SEMP). The present sequence is one such probe. The SEMP are derived

CC from human HeLa cells. The SEMP can be used to produce a single exon

CC microarray, which can be used for measuring human gene expression in a

CC sample derived from human cervical epithelial cells. By measuring gene

CC expression, the probes are therefore useful in grading and/or staging

CC of diseases of the cervix, notably cervical cancer.

CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

XX at ftp.wipo.int/pub/published_pcl_sequences.

XX Sequence 294 BP; 69 A; 60 C; 79 G; 86 T; 0 other;

XX Alignment Scores:

Alignment Scores:	0.0708	Length:	294
Pred. No.:	12.00	Matches:	12
Score:	100.00%	Conservative:	0
Percent Similarity:	100.00%	Mismatches:	0
Best Local Similarity:	100.00%	Indels:	0
Query Match:	1.09%	Gaps:	0
DB:	22		


```

US-09-397-967-16 (1-1099) x AA123695 (1-294)
OY 941 ValHisArgAspLeuAlaIaArgAsnIleLeuVal 952
   |||||||
DB 91 GTGCATCGTATCTGCCCGACGACATCTCTG 126

RESULT 32
AA149009
ID AA149009 standard: DNA; 294 BP.
XX
AC AA149009;
XX
DT 17-OCT-2001 (first entry)
XX
XX
DE Probe #17695 used to measure gene expression in human placenta sample.
XX
XX
KM Probe; microarray; human; placenta; antenatal diagnosis;
KM genetic disorder; ss.
OS Homo sapiens.
XX
PN WO200157272-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00663.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488897/53.
XX
XX
PT Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human placenta -
XX
PS Claim 25; SEQ ID No 17695; 654bp; English.
XX
CC The present invention relates to single exon nucleic acid probes (SENP).
CC The present sequence is one such probe. The probes are useful for
CC producing a microarray for predicting, measuring and displaying gene
CC expression in samples derived from human placenta. The probes are useful
CC for antenatal diagnosis of human genetic disorders.
XX
SQ Sequence 294 BP; 69 A; 60 C; 79 G; 86 T; 0 other;
XX

Alignment Scores:
Pred. No.: 0.0708 Length: 294
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 22 Gaps: 0

US-09-397-967-16 (1-1099) x AA149009 (1-294)
OY 941 ValHisArgAspLeuAlaIaArgAsnIleLeuVal 952
   |||||||
DB 91 GTGCATCGTATCTGCCCGACGACATCTCTG 126

RESULT 33
AA109314
ID AA109314 standard: DNA; 294 BP.
XX
AC AA109314;
XX

```

```

XX
DT 09-OCT-2001 (first entry)
XX
XX
DE Probe #9305 used to measure gene expression in human breast sample.
XX
XX
KM Probe; human; breast disease; breast cancer; development disorder; ss;
KM inflammatory disease; proliferative breast disease; non-carcinoma tumour.
OS Homo sapiens.
XX
PN WO200157270-A2.
XX
PD 09-AUG-2001.
XX
PF 29-JAN-2001; 2001WO-US00661.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-476286/51.
XX
XX
PT Novel single exon nucleic acid probe used to measuring gene expression
XX in a human breast -
XX
PS Claim 25; SEQ ID No 9305; 322pp; English.
XX
XX
CC The present invention relates to novel single exon nucleic acid probes.
CC The present sequence is one such probe. The probes are useful for
CC measuring human gene expression in a human breast sample, where the probe
CC hybridises at high stringency to a nucleic acid expressed in the human
CC breast. The probes are useful for predicting, diagnosing, grading,
CC staging, monitoring and prognosing diseases of the human breast,
CC particularly those diseases with polygenic aetiology. The diseases
CC include: breast cancer; disorders of development, inflammatory diseases
CC of the breast; fibrocystic changes; proliferative breast disease and
CC non-carcinoma tumours.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published-pct-sequences.
XX
SQ Sequence 294 BP; 69 A; 60 C; 79 G; 86 T; 0 other;
XX

Alignment Scores:
Pred. No.: 0.0708 Length: 294
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 22 Gaps: 0

US-09-397-967-16 (1-1099) x AA109314 (1-294)
OY 941 ValHisArgAspLeuAlaIaArgAsnIleLeuVal 952
   |||||||
DB 91 GTGCATCGTATCTGCCCGACGACATCTCTG 126

RESULT 34
ABS16994
ID ABS16994 standard: DNA; 294 BP.
XX
AC ABS16994;
XX
DT 19-AUG-2002 (first entry)
XX

```

DE Human genome-derived single exon probe ORF from lung SEQ ID No 16985.
 XX Human: ds; single exon probe; asthma; lung cancer; COPD; ILD;
 KW chronic obstructive pulmonary disease; interstitial lung disease;
 KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
 KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
 KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemorrhoidosis;
 KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
 KW pulmonary alveolar proteinosis; fibrocytic pulmonary dysplasia;
 KW primary ciliary dyskinesia; pulmonary hypertension;
 KW hyaline membrane disease; open reading frame; ORF.
 XX Homo sapiens.
 XX MO200186003-A2.
 PN 15-NOV-2001.
 PD 30-JAN-2001; 2001WO-US00665.
 PF 04-FEB-2000; 2000US-180312P.
 PR 26-MAY-2000; 2000US-207456P.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-234687P.
 PR 27-SEP-2000; 2000US-236359P.
 PR 04-OCT-2000; 2000GB-0024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 PA Penn SG, Hanzel DK, Chen W, Rank DR;
 PI WPI; 2002-114183/15.
 DR Spatially-addressable set of single exon nucleic acid probes, used to
 PT measure gene expression in human lung samples -
 PS Claim 4; SEQ ID No 16985; 634pp; English.
 XX The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human lung comprising single exon nucleic acid probes having one of
 CC 12614 nucleic acid sequences mentioned in the specification, or their
 CC complements or the 12387 open reading frames derived from the 12614
 CC probes. Also included are a microarray comprising the novel set of
 CC probes. The novel set of probes which hybridize at high stringency to a
 CC nucleic acid expressed in the human lung; measuring gene expression in a
 CC sample derived from human lung, comprising (a) contacting the array with
 CC a collection of detectably labeled nucleic acids derived from human lung
 CC mRNA, and (b) measuring the label detectably bound to each probe of
 CC the array; identifying exons in a eukaryotic genome, comprising
 CC (a) algorithmically predicting at least one exon from genomic sequences
 CC of the eukaryote; and (b) detecting specific hybridization of detectably
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
 CC having a fragment identical to the predicted exon, the probe is included
 CC in the above mentioned microarray; assigning exons to a single gene,
 CC comprising (a) identifying exons from genomic sequence by the method
 CC above and (b) measuring the expression of each of the exons in several
 CC tissues and/or cell types using hybridization to a single exon
 CC microarrays having a probe with the exon, where a common pattern of
 CC expression of the exons in the tissues and/or cell types indicates that
 CC the exons should be assigned to a single gene; a peptide comprising one
 CC of 12011 sequences, mentioned in the specification, or encoded by the
 CC probes/open reading frames (ORF). The probes are used for gene
 CC expression analysis, and for identifying exons in a gene, particularly
 CC using human lung derived mRNA and for the study of lung diseases
 CC such as asthma, lung cancer, chronic obstructive pulmonary disease
 CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
 CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
 CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
 CC haemorrhoidosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
 CC pulmonary alveolar proteinosis, Karagener syndrome, fibrocytic
 CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension

CC and hyaline membrane disease. The present sequence is a single exon
 CC probe open reading frame of the invention.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 294 BP: 69 A; 60 C; 79 G; 86 T; 0 other;
 Alignment Scores:
 Pred. No.: 0.0708 Length: 294
 Score: 12.00 Matches: 12
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.09% Indels: 0
 DB: 24 Gaps: 0
 US-09-397-967-16 (1-1099) x ABS16994 (1-294)
 QY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
 |||||||||||||||||||||||||||||||||||
 Db 91 GTGCATCTGATCTGCGCCGACGACATCTCTGTG 126
 RESULT 35
 ABA45718
 ID ABA45718 standard; DNA; 367 BP.
 XX
 AC ABA45718;
 XX
 DT 01-FEB-2002 (first entry)
 DE
 DE Human breast cell single exon nucleic acid probe #4413.
 XX
 KW Human: microarray; single exon probe; gene expression; breast;
 KW disease; cancer; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200157271-A2.
 XX
 PD 09-AUG-2001.
 PF 30-JAN-2001; 2001WO-US00662.
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 PA Penn SG, Hanzel DK, Chen W, Rank DR;
 PI WPI; 2001-496933/54.
 DR New spatially-addressable set of single exon nucleic acid probes,
 PT useful for measuring gene expression in sample derived from human
 PT breast, comprises number of single exon nucleic acid probes -
 PS Claim 1; SEQ ID NO 4413; 327pp + sequence listing; English.
 XX The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human breast and BT 474 cells. The method involves contacting
 CC the probes with a collection of detectably labeled nucleic acids
 CC derived from mRNA of human breast, and then measuring the label
 CC bound to each probe of the microarray. The probes are useful for
 CC verifying the expression of regions of genomic DNA predicted to
 CC encode proteins. They are useful for gene discovery, and for
 CC determining predisposition and/or prognosing breast disease. Gene

CC expression analysis is useful for assessing the toxicity of chemical
 CC agents on cells. The microarray of this invention presents a far greater
 CC diversity of probes for measuring gene expression, with far less bias
 CC than expressed sequence tag microarrays. The method is suitable for
 CC rapid production of functional information from genomic sequence. The
 CC present sequence is a single exon nucleic acid probe of the invention.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 CC

Sequence 367 BP; 102 A; 71 C; 89 G; 105 T; 0 other;

Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
Score:	0.0872	367	12	0	0	0
Percent Similarity:	12.00					
Best Local Similarity:	100.00%					
Query Match:	1.09%					
DB:	22					

US-09-397-967-16 (1-1099) x ABA5718 (1-367)

OY 941 ValHisArgAspLeuAlaAlaArgAsnIleuVal 952
 Db 289 GTGCATCGTATCTGCGCCGACGAAACATCTGCTG 324

RESULT 36

ABAS6224
 ID ABA56224 standard; DNA: 367 BP.

XX ABA56224;

DT 01-FEB-2002 (first entry)

DE Human foetal liver single exon nucleic acid probe #4529.

XX Human; foetal liver; gene expression; single exon nucleic acid probe; ss.

OS Homo sapiens.

PN WO200157277-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00669.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PA Penn SG, Hanzel DK, Chen W, Rank DR;

PI WPI; 2001-483447/52.

PT Human genome-derived single exon nucleic acid probes useful for

XX analyzing gene expression in human foetal liver -

PS Claim 1; SEQ ID NO 4529; 639bp + sequence listing; English.

CC The invention relates to a single exon nucleic acid probe for
 CC measuring human gene expression in a sample derived from human foetal
 CC liver. The single exon nucleic acid probes may be used for predicting,
 CC measuring and displaying gene expression in samples derived from human
 CC foetal liver. The present sequence is a single exon nucleic acid
 CC probe of the invention.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly

CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 367 BP; 102 A; 71 C; 89 G; 105 T; 0 other;

Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
Score:	0.0872	367	12	0	0	0
Percent Similarity:	12.00					
Best Local Similarity:	100.00%					
Query Match:	1.09%					
DB:	22					

US-09-397-967-16 (1-1099) x ABA56224 (1-367)

OY 941 ValHisArgAspLeuAlaAlaArgAsnIleuVal 952
 Db 289 GTGCATCGTATCTGCGCCGACGAAACATCTGCTG 324

RESULT 37

ABA25869
 ID ABA25869 standard; DNA: 367 BP.

XX ABA25869;

DT 23-JAN-2002 (first entry)

DE Probe #4335 for gene expression analysis in human heart cell sample.

XX Human; gene expression; heart; microarray; vascular system; probe;

KW cardiovascular disease; hypertension; cardiac arrhythmia;

XX congenital heart disease; ss.

OS Homo sapiens.

PN WO200157274-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00666.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PA Penn SG, Hanzel DK, Chen W, Rank DR;

PI WPI; 2001-488899/53.

PT Single exon nucleic acid probes for analyzing gene expression in human

XX hearts -

PS Claim 1; SEQ ID NO 4335; 530bp; English.

CC The present invention relates to single exon nucleic acid probes for
 CC measuring human gene expression in a sample derived from human heart. The
 CC present sequence is one such probe. The probes may be used for
 CC predicting, measuring and displaying gene expression in samples derived
 CC from the human heart via microarrays. By measuring gene expression, the
 CC probes are useful for predicting, diagnosing, grading, staging,
 CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
 CC congenital heart disease.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC

Sequence 367 BP; 102 A; 71 C; 89 G; 105 T; 0 other;

Alignment Scores:

Pred. No.:	0.0872	Length:	367
Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.09%	Indels:	0
DB:	22	Gaps:	0

US-09-397-967-16 (1-1099) x ABA25869 (1-367)

OY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
 |||||
 Db 289 GTGCATCGTGCATCTGCGCCGACGAAACATCTCGTGG 324

RESULT 38
 AAK04412
 ID AAK04412 standard; DNA; 367 BP.

AC AAK04412;
 DT 05-NOV-2001 (first entry)
 XX
 DE Human brain expressed single exon probe SEQ ID NO: 4403.
 XX
 KW Human; brain expressed exon; gene expression analysis; probe;
 KM microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
 KW epilepsy; cancer; ss.
 OS Homo sapiens.
 XX
 PN WO200157275-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00667.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 PR
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-483446/52.
 XX
 DR Single exon nucleic acid probes for analyzing gene expression in human
 PT brains -
 PT
 PS Example 4; SEQ ID NO: 4403; 650pp + Sequence Listing; English.
 XX
 CC The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC brain. They can be used to measure gene expression in brain cell samples,
 CC which may enable the diagnosis and improved treatment of nervous system
 CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
 CC epilepsy and cancers. The present sequence is one of the probes of the
 CC invention.
 CC
 XX
 SQ Sequence 367 BP; 102 A; 71 C; 89 G; 105 T; 0 other;

Alignment Scores:

Pred. No.:	0.0872	Length:	367
Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.09%	Indels:	0
DB:	22	Gaps:	0

US-09-397-967-16 (1-1099) x AAK04412 (1-367)

OY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
 |||||
 Db 289 GTGCATCGTGCATCTGCGCCGACGAAACATCTCGTGG 324

RESULT 39
 AAK29908
 ID AAK29908 standard; DNA; 367 BP.

AC AAK29908;
 DT 06-NOV-2001 (first entry)
 XX
 DE Human bone marrow expressed single exon probe SEQ ID NO: 4465.
 XX
 KW Human; bone marrow expressed exon; gene expression analysis; probe;
 KM microarray; cancer; leukemia; lymphoma; myeloma; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200157276-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00668.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 PR
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-488900/53.
 XX
 DR Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human bone marrow -
 PT
 PS Example 4; SEQ ID NO: 4465; 658pp + Sequence Listing; English.
 XX
 CC The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC bone marrow. They can be used to measure gene expression in bone marrow
 CC samples, which may enable the improved diagnosis and treatment of cancers
 CC such as lymphoma, leukemia and myeloma. The present sequence is one of
 CC the probes of the invention.
 CC
 XX
 SQ Sequence 367 BP; 102 A; 71 C; 89 G; 105 T; 0 other;

Alignment Scores:

Pred. No.:	0.0872	Length:	367
Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.09%	Indels:	0
DB:	22	Gaps:	0

US-09-397-967-16 (1-1099) x AAK29908 (1-367)

OY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
 |||||
 Db 289 GTGCATCGTGCATCTGCGCCGACGAAACATCTCGTGG 324

RESULT 40
 AAI14495
 ID AAI14495 standard; DNA; 367 BP.

XX AC AAI14495;
XX KW Probe: microarray; human; placenta; antenatal diagnosis;
XX KW genetic disorder; ss.
XX OS Homo sapiens.
XX PN WO200157272-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US006670.
XX PR 04-FEB-2000; 2000US-0180312.
XX PR 26-MAY-2000; 2000US-0207456.
XX PR 30-JUN-2000; 2000US-0608408.
XX PR 03-AUG-2000; 2000US-0632366.
XX PR 21-SEP-2000; 2000US-0234687.
XX PR 27-SEP-2000; 2000US-0236359.
XX PR 04-OCT-2000; 2000GB-0024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX DR WPI: 2001-488901/53.
XX PS Claim 25; SEQ ID No 4428; 487pp; English.
XX CC The present invention relates to human single exon nucleic acid probes (SNP). The present sequence is one such probe. The SNPs are derived from human HeLa cells. The SNPs can be used to produce a single exon microarray, which can be used for measuring human gene expression in a sample derived from human cervical epithelial cells. By measuring gene expression, the probes are therefore useful in grading and/or staging gene of diseases of the cervix, notably cervical cancer.
XX CC Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX CC
XX CC
XX SQ Sequence 367 BP; 102 A; 71 C; 89 G; 105 T; 0 other;
XX
XX Alignment Scores:
XX Pred. No.: 0.0872 Length: 367
XX Score: 12.00 Matches: 12
XX Percent Similarity: 100.00% Conservative: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 1.09% Indels: 0
XX DB: 22 Gaps: 0
XX
XX US-09-397-967-16 (1-1099) x AAI14495 (1-367)
XX QY 941 ValHsArgAspLeuAlaAlaArgAsnIleLeuVal 952
XX Db 289 GTGCATCTGATCTGGCCGACGGAACATCTCTGGTG 324
XX
XX RESULT 41
XX ID AAI35871 standard; DNA; 367 BP.
XX AC AAI35871;
XX XX
XX DT 17-OCT-2001 (first entry)
XX XX
XX DE Probe #4557 used to measure gene expression in human placenta sample.

XX AC AAI14495;
XX KW Probe: microarray; human; placenta; antenatal diagnosis;
XX KW genetic disorder; ss.
XX OS Homo sapiens.
XX PN WO200157272-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US006670.
XX PR 04-FEB-2000; 2000US-0180312.
XX PR 26-MAY-2000; 2000US-0207456.
XX PR 30-JUN-2000; 2000US-0608408.
XX PR 03-AUG-2000; 2000US-0632366.
XX PR 21-SEP-2000; 2000US-0234687.
XX PR 27-SEP-2000; 2000US-0236359.
XX PR 04-OCT-2000; 2000GB-0024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX DR WPI: 2001-488901/53.
XX PS Claim 25; SEQ ID No 4557; 654pp; English.
XX CC The present invention relates to single exon nucleic acid probes (SNP). The present sequence is one such probe. The probes are useful for producing a microarray for predicting, measuring and displaying gene expression in samples derived from human placenta. The probes are useful for antenatal diagnosis of human genetic disorders.
XX CC
XX CC
XX SQ Sequence 367 BP; 102 A; 71 C; 89 G; 105 T; 0 other;
XX
XX Alignment Scores:
XX Pred. No.: 0.0872 Length: 367
XX Score: 12.00 Matches: 12
XX Percent Similarity: 100.00% Conservative: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 1.09% Indels: 0
XX DB: 22 Gaps: 0
XX
XX US-09-397-967-16 (1-1099) x AAI35871 (1-367)
XX QY 941 ValHsArgAspLeuAlaAlaArgAsnIleLeuVal 952
XX Db 289 GTGCATCTGATCTGGCCGACGGAACATCTCTGGTG 324
XX
XX RESULT 42
XX ID AAI04322 standard; DNA; 367 BP.
XX AC AAI04322;
XX XX
XX DT 09-OCT-2001 (first entry)
XX XX
XX DE Probe #4313 used to measure gene expression in human breast sample.
XX KW Probe: human; breast disease; breast cancer; development disorder; ss;
XX KW inflammatory disease; proliferative breast disease; non-carcinoma tumour.
XX OS Homo sapiens.
XX PN WO200157270-A2.
XX PD 09-AUG-2001.
XX PF 29-JAN-2001; 2001WO-US00661.

```

XX 04-FEB-2000: 2000US-0180312.
PR 26-MAY-2000: 2000US-0207456.
PR 30-JUN-2000: 2000US-0608408.
PR 03-AUG-2000: 2000US-0632366.
PR 21-SEP-2000: 2000US-0234687.
PR 27-SEP-2000: 2000US-0236359.
PR 04-OCT-2000: 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR:
XX WPI: 2001-476286/51.
XX
XX Novel single exon nucleic acid probe used to measuring gene expression
XX in a human breast -
XX
XX Claim 25: SEQ ID NO 4313; 322pp; English.
XX
XX The present invention relates to novel single exon nucleic acid probes.
XX The present sequence is one such probe. The probes are useful for
XX measuring human gene expression in a human breast sample, where the probe
XX hybridizes at high stringency to a nucleic acid expressed in the human
XX breast. The probes are useful for predicting, diagnosing, grading,
XX staging, monitoring and prognosing diseases of the human breast,
XX particularly those diseases with polygenic aetiology. The diseases
XX include: breast cancer, disorders of development, inflammatory diseases
XX of the breast, fibrocystic changes, proliferative breast disease and
XX non-carcinoma tumours.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 367 BP: 102 A; 71 C; 89 G; 105 T; 0 other:

Alignment Scores:
Pred. No.: 0.0872 Length: 367
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: Gaps: 0

US-09-397-967-16 (1-1099) x AA104322 (1-367)
OY 941 ValHisArgAspLeuAlaAlaArgAsnIleuVal 952
DB 289 GTGCATCGTGAATGCGCCGACGGACATCTGTG 324

RESULT 43
ABS04474
ID ABS04474 standard; DNA; 367 BP.
XX
XX ABS04474:
XX
XX 19-AUG-2002 (first entry)
XX
XX Human genome-derived single exon probe from lung SEQ ID No 4465.
XX
XX Human: ds; single exon probe; asthma; lung cancer; COPD; IHD;
XX chronic obstructive pulmonary disease; interstitial lung disease;
XX familial idiopathic pulmonary fibrosis; neurofibromatosis;
XX tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
XX Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
XX pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
XX pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
XX primary ciliary dyskinesia; pulmonary hypertension;
XX hyaline membrane disease.
XX
XX Homo sapiens.
XX
XX WO200186003-A2.

```

```

XX 15-NOV-2001.
PD
XX
XX 30-JAN-2001: 2001WO-US00665.
PF
XX
XX 04-FEB-2000: 2000US-180312P.
PR 26-MAY-2000: 2000US-207456P.
PR 30-JUN-2000: 2000US-0608408.
PR 03-AUG-2000: 2000US-0632366.
PR 21-SEP-2000: 2000US-234687P.
PR 27-SEP-2000: 2000US-236359P.
PR 04-OCT-2000: 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR:
XX WPI: 2002-114183/15.
XX
XX Spatially-addressable set of single exon nucleic acid probes, used to
XX measure gene expression in human lung samples -
XX
XX Claim 1: SEQ ID NO 4465; 634pp; English.
XX
XX The invention relates to a spatially-addressable set of single exon
XX nucleic acid probes for measuring gene expression in a sample derived
XX from human lung comprising single exon nucleic acid probes having one of
XX 12614 nucleic acid sequences mentioned in the specification, or their
XX complements or the 12387 open reading frames derived from the 12614
XX probes. Also included are a microarray comprising the novel set of
XX probes; the novel set of probes which hybridize at high stringency to a
XX nucleic acid expressed in the human lung; measuring gene expression in a
XX sample derived from human lung, comprising (a) contacting the array with
XX a collection of detectably labeled nucleic acids derived from human lung
XX mRNA, and (b) measuring the label detectably bound to each probe of
XX the array; identifying exons in a eukaryotic genome, comprising
XX (a) algorithmically predicting at least one exon from genomic sequences
XX of the eukaryote; and (b) detecting specific hybridisation of detectably
XX labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
XX having a fragment identical to the predicted exon, the probe is included
XX in the above mentioned microarray; assigning exons to a single gene,
XX comprising (a) identifying exons from genomic sequence by the method
XX above and (b) measuring the expression of each of the exons in several
XX tissues and/or cell types using hybridisation to a single exon
XX microarrays having a probe with the exon, where a common pattern of
XX expression of the exons in the tissues and/or cell types indicates that
XX the exons should be assigned to a single gene; a peptide comprising one
XX of 12011 sequences, mentioned in the specification, or encoded by the
XX probes/open reading frames (ORF). The probes are used for gene
XX expression analysis, and for identifying exons in a gene, particularly
XX using human lung derived mRNA and for the study of lung diseases
XX such as asthma, lung cancer, chronic obstructive pulmonary disease
XX (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
XX fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
XX Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
XX haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
XX pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic
XX pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
XX and hyaline membrane disease. The present sequence is a single exon
XX probe of the invention.
XX Note: The sequence data for this patent did not form part
XX of the printed specification, but was obtained in electronic
XX format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 367 BP: 102 A; 71 C; 89 G; 105 T; 0 other:

Alignment Scores:
Pred. No.: 0.0872 Length: 367
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0

```


CC used as diagnostic agents for detecting the presence of flea polypeptides
 CC in samples (e.g., by enzyme linked immunosorbent assay (ELISA)). The
 CC present sequence represents a cat flea HMT cDNA of the invention.

SO Sequence 917 BP; 263 A; 175 C; 207 G; 242 T; 30 other;

Alignment Scores:

Pred. No.:	0.206	Length:	917
Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.09%	Indels:	0
DB:	21	Gaps:	0

US-09-397-967-16 (1-1099) x AAC95082 (1-917)

OY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952

DB 169 GTCCATCGAGATTGGCTGCCAGAAATATCTGTG 204

RESULT 46

ID AAS87117 standard; cDNA; 1117 BP.

XX AAS87117:

DT 13-FEB-2002 (first entry)

DE DNA encoding novel human diagnostic protein #22921.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX Homo sapiens.

XX WO200175067-A2.

PD 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US08631.

XX 31-MAR-2000; 2000US-0540217.

XX 23-AUG-2000; 2000US-0649167.

PA (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

DR WPI; 2001-639362/73.

DR P-PSDB; AAG22930.

PT New isolated polynucleotide and encoded polypeptides, useful in

PT diagnostics, forensics, gene mapping, identification of mutations

PT responsible for genetic disorders or other traits and to assess

PT biodiversity

PS Claim 1; SEQ ID NO 22921; 103pp; English.

CC The invention relates to isolated polynucleotide (I) and

CC polypeptide (II) sequences. (I) is useful as hybridisation probes,

CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome

CC and gene mapping, and in recombinant production of (II). The

CC polynucleotides are also used in diagnostics as expressed sequence tags

CC for identifying expressed genes. (I) is useful in gene therapy techniques

CC to restore normal activity of (II) or to treat disease states involving

CC (II). (II) is useful for generating antibodies against it, detecting or

CC quantitating a polypeptide in tissue, as molecular weight markers and as

CC a food supplement. (II) and its binding partners are useful for treating

CC disorders involving aberrant protein expression or biological activity.

CC The polypeptide and polynucleotide sequences have applications in

CC diagnostics, forensics, gene mapping, identification of mutations

CC responsible for genetic disorders or other traits to assess biodiversity

CC and to produce other types of data and products dependent on DNA and

CC amino acid sequences. AAS64197-AAS94564 represent novel human

CC diagnostic coding sequences of the invention.

CC Note: The sequence data for this patent did not appear in the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pcl_sequences.

SO Sequence 1117 BP; 308 A; 246 C; 292 G; 271 T; 0 other;

Alignment Scores:

Pred. No.:	0.248	Length:	1117
Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.09%	Indels:	0
DB:	23	Gaps:	0

US-09-397-967-16 (1-1099) x AAS87117 (1-1117)

OY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952

DB 235 GTCCATCGATGATCTGCCAGCAACATCTGTG 270

RESULT 47

ID AAO53470 standard; DNA; 1509 BP.

XX AAO53470:

DT 16-JUN-1994 (first entry)

DE PKR gene B1.

KW Lambda gt11; expression vector; lambda-B1-Elk; protein tyrosine kinase;

KW Elk; B1; Eph; subfamily; receptor-like tyrosine kinase; eph; eck;

KW phosphorylation; phosphorylated kinase insert domain; growth factor;

KW receptor kinase; platelet-derived growth factor receptor; ss.

XX Rattus rattus.

XX CA2083521-A.

XX 01-OCT-1993.

XX 23-NOV-1992; 92CA-2083521.

XX 31-MAR-1992; 92US-0861390.

PA (MOUN) MOUNT SINAI HOSPITAL CORP.

PI Letwin K, Pawson A, Reedijk M;

DR WPI; 1993-406300/51.

DR P-PSDB; AAR44512.

PT Expression of phosphorylated exogenous protein - in host cells

PT transformed with two vectors, one for the protein, the other for

PT catalytic domain of protein kinase

PS Disclosure; Fig 1; 55pp; English.

CC This sequence represents a fragment of the lambda gt11 expression

CC vector, lambda-B1-Elk, which encodes the catalytic sequence of the

CC protein tyrosine kinase, Elk. The Elk gene, B1, encodes a protein

CC which is a member of the Eph subfamily of protein tyrosine kinases.

CC The Elk product is very similar to two other receptor-like tyrosine

CC kinases, eph and eck. Lambda-B1-Elk may be used in the production

CC of phosphorylated exogenous protein along with a further vector

CC encoding the desired exogenous protein. These plasmid may be used
 CC to produce phosphorylated proteins in host cells which have no
 CC intrinsic capacity for phosphorylation, eg. bacteria. The system
 CC may be used for the expression of the phosphorylated kinase insert
 CC domain of a growth factor receptor kinase eg. platelet-derived growth
 CC factor receptor.

XX Sequence 1509 BP; 393 A; 374 C; 429 G; 313 T; 0 other;

Alignment Scores:

Pred. No.:	0.329	Length:	1509
Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.09%	Indels:	0
DB:	14	Gaps:	0

US-09-397-967-16 (1-1099) x AAO53470 (1-1509)

OY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952

DB 409 GTGCACCGGACCTGCGCTAGCAACATCTCTGCTG 444

RESULT 48

AAO84888

ID AAO84888 standard; cDNA to mRNA; 1942 BP.

AC AAO84888;

XX 26-OCT-1995 (first entry)

DE DNA encoding cytoplasmic tyrosine kinase.

KW cytoplasmic; tyrosine kinase; blood; cell differentiation;

KW screening; anticancer agent; ds.

XX Homo sapiens.

XX OS

XX

Location/Qualifiers

FT Key

FT CDS

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

/note="encodes N-terminaly truncated form
 of the enzyme (see AAR71132)."

XX W09506113-A.

XX 02-MAR-1995.

XX 25-AUG-1994;

XX 94WO-JP01411.

XX 25-AUG-1993;

XX 93JP-0210403.

XX 29-MAR-1994;

XX 94JP-0058553.

XX (ASAH) ASAMI KASEI KOGYO KK.

XX Sakano S;

XX WPI: 1995-106842/14.

XX P-PSDB; AAR71129-33.

XX Cytoplasmic tyrosine kinase and antibody recognising it - for
 PT screening chemical substances for tyrosine kinase inhibitory or
 PT activating activity for use as cancer therapy

PS Claim 7; Page 49-50; 58pp; English.

CC This DNA encodes a cytoplasmic tyrosine kinase which has enhanced
 CC expression in connection with blood cell differentiation. It was
 CC isolated from the human HT-7 blood cell line. The DNA sequences and
 CC antibodies raised against the enzyme, are useful for screening agents
 CC for inhibiting or activating activity on the tyrosine kinase, for
 CC use as anticancer agents.

XX Sequence 1942 BP; 365 A; 615 C; 651 G; 311 T; 0 other;

Alignment Scores:

Pred. No.:	0.416	Length:	1942
Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.09%	Indels:	0
DB:	16	Gaps:	0

US-09-397-967-16 (1-1099) x AAO84888 (1-1942)

OY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952

DB 1252 GTGCACCGGACCTGCGCGCCGCAACATCTGCTG 1287

RESULT 49

AAV44497

ID AAV44497 standard; cDNA; 1987 BP.

AC AAV44497;

XX 16-OCT-1998 (first entry)

DE Human matk cDNA.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

/note="CSK homologous kinase"

XX W09830704-A1.

XX 16-JUL-1998.

XX 07-JAN-1998;

XX 98WO-US00420.

XX 16-JUN-1997;

XX 97US-0876882.

XX 08-JAN-1997;

XX 97US-0035228.

XX (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.

XX Avraham H. Groopman JE;

XX WPI: 1998-399149/34.

XX P-PSDB; AAV44454.

XX

XX

XX

XX

XX

XX

Detecting breast cancer by detecting Csk homologous kinase
 expression - especially in humans and use of Csk homologous kinase
 in treatment or prophylaxis of breast cancer and for producing
 medicaments

PS Disclosure; Fig 2; 54pp; English.

XX This sequence encodes a CKK homologous kinase (CHK) which is used in a
 CC method of detecting cancer in breast tissue. The method allows diagnosis
 CC of breast cancer in mammals, especially humans. It is based on the
 CC discovery that a cytoplasmic protein tyrosine kinase, CHK, is expressed
 CC in human breast tissue, but not in adjacent tissue. This protein can be
 CC used to raise antibodies which can be included in compositions and
 CC diagnostic kits for diagnosis of breast cancer. The presence of CHK in
 CC breast tissue can also be determined using other standard methods (e.g.,
 CC Northern blotting) or by detecting nucleic acid sequences encoding all/a
 CC portion of the protein (e.g., using hybridisation probes). Over-expression
 CC of the receptor tyrosine kinase ErbB-2 has previously been associated
 CC with the development of breast cancer, and CHK specifically interacts
 CC with activated ErbB-2, and may function as a negative regulator of
 CC ErbB-2 mediated mitogenic signalling. The compositions may also be used
 CC to design drugs (e.g., which incorporate CHK analogues with greater
 CC biological activity than CHK) and to identify CHK antagonists and
 CC agonists for therapeutic use.

XX Sequence 1987 BP; 369 A; 628 C; 672 G; 318 T; 0 other;

SQ Alignment Scores:

Pred. No.:	0.425	Length:	1987
Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.09%	Indels:	0
DB:	19	Gaps:	0

US-09-397-967-16 (1-1099) x AAV44497 (1-1987)

OY 941 VALHISARGASPLEUALAALARGASNILEUVAL 952
 |||||||||||||||||||||||||||||||||||

DB 1307 GTGCACCGCGACCTGGCGCGCACATCTGTGTC 1342

RESULT 50
 AAT00616
 ID AAT00616 standard; cDNA; 2000 BP.

XX AAT00616;

AC AAT00616;

XX 26-MAR-1996 (first entry)

DT Megakaryocyte kinase MKK1 cDNA.

XX Megakaryocyte kinase MKK1; cytoplasmic tyrosine kinase;
 KW cellular signal transduction; leukaemia; myelosis; myelofibrosis;
 KW gene therapy; ss.

XX Homo sapiens.

OS Homo sapiens.

XX Key Location/Qualifiers

FT CDS 258..1781

FT /tag- a

XX MO9529185-A1.

XX 02-NOV-1995.

PD 24-APR-1995; 95MO-US05008.

XX 21-APR-1995; 95US-0426509.

PR 22-APR-1994; 94US-0232545.

XX (PLAC) MAX PLANCK GES FÖRDERUNG WISSENSCHAFTEN.
 PA (SUG-) SUGEN INC.

XX Gishizky M, Sures I., Ullrich A;
 PI WPL: 1995-382959/49.
 DR P-PSDB: AAR84181.
 XX

PT New polynucleotide(s) encoding megakaryocyte tyrosine kinase(s) -
 PT used to develop prods. for the treatment and diagnosis of kinase
 PT related signal transduction abnormalities.

XX Claim 2; Fig 1A-C; 82pp; English.

PS Overlapping cDNA clones were combined to obtain a sequence (AAT00616)
 CC coding for human megakaryocyte kinase MKK1 (AAR84181). The clones
 CC were isolated from a fetal brain library using degenerate primers
 CC (AAT00614-15) based on conserved regions within the kinase domain of
 CC receptor tyrosine kinases. MKK polynucleotides can be used in the
 CC prodn. in host cells of recombinant MKK, and in the gene therapy of
 CC diseases such as acute megakaryocytic leukaemia, myelofibrosis and
 CC acute megakaryocytic myelosis.

XX Sequence 2000 BP; 375 A; 631 C; 674 G; 320 T; 0 other;

SQ Alignment Scores:

Pred. No.:	0.428	Length:	2000
Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.09%	Indels:	0
DB:	16	Gaps:	0

US-09-397-967-16 (1-1099) x AAT00616 (1-2000)

OY 941 VALHISARGASPLEUALAALARGASNILEUVAL 952
 |||||||||||||||||||||||||||||||||||

DB 1302 GTGCACCGCGACCTGGCGCGCACATCTGTGTC 1337

RESULT 51
 AA241307
 ID AA241307 standard; cDNA; 2170 BP.

XX AA241307;

AC AA241307;

XX 18-JAN-2000 (first entry)

DT Human normal ovarian tissue derived cDNA 86.

XX Human; ovary; screening; ovarian cancer; treatment; ss.

XX Homo sapiens.

OS Homo sapiens.

XX DE19816395-A1.

PN 07-OCT-1999.

XX 03-APR-1998; 98DE-1016395.

PF 03-APR-1998; 98DE-1016395.

XX (META-) METAGEN GES GENOMFORSCHUNG MBH.

PA Rösenthal A, Specht T, Hinzmann B, Schmitt A, Piliarsky C, Dahl E;
 PI WPL: 1999-552352/47.

DR Nucleic acid sequences potentially useful in diagnosis or therapy of
 PT ovarian cancer

XX Claim 3; Page 196; 274pp; German.

XX This invention describes novel nucleic acid sequences that are highly
 CC expressed in normal ovary tissue. Artificial chromosomes and cosmid
 CC clones containing the sequences can be used as gene transfer vehicles.
 CC The sequences can be used to produce DNA fragments containing
 CC full-length genes. Host cells transformed with the sequences can be used
 CC to produce polypeptides or polypeptide fragments, which can be used to
 CC screen phage displays for polypeptides that bind to them, or as tools for
 CC identifying agents active against ovarian cancer, or to prepare
 CC medicaments for treating ovarian cancer. The cDNA sequences can be used

CC to obtain genomic genes, their promoters, enhancers, silencers, exon
CC structures, intron structures and their splice variants. AA241222-241324
CC represent cDNA sequences derived from normal human ovarian tissue and
CC which encode the protein fragments represented in AAY59724-Y59837.
XX

SO Sequence 2170 BP; 483 A; 605 C; 615 G; 467 T; 0 other;

Alignment Scores:

Pred. No.:	0.462	Length:	2170
Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.09%	Indels:	0
DB:	20	Gaps:	0

US-09-397-967-16 (1-1099) x AA241307 (1-2170)

OY 941 VALHISARGASPLEUALAALARGASNTLEUVAL 952

DB 485 GTTCACCGAGACTGCTGCTGCACACATCTAGTC 520

RESULT 52

AAS85547
ID AAS85547 standard; CDNA: 2489 BP.

XX AAS85547;

XX 13-FEB-2002 (first entry)

XX DNA encoding novel human diagnostic protein #21351.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;

XX Food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX Homo sapiens.

XX MO200125067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US08631.

XX 31-MAR-2000; 2000US-0540217.

XX 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI: 2001-639362/73.

XX P-PSDB; ABG21360.

XX New isolated polynucleotide and encoded polypeptides, useful in

XX diagnostics, forensics, gene mapping, identification of mutations

XX bioldiversity

XX Claim 1; SEQ ID NO 21351; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and

XX polypeptide (II) sequences. (I) is useful as hybridisation probes,

XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome

XX and gene mapping, and in recombinant production of (II). The

XX polynucleotides are also used in diagnostics as expressed sequence tags

XX for identifying expressed genes. (I) is useful in gene therapy techniques

XX to restore normal activity of (II) or to treat disease states involving

XX (II). (II) is useful for generating antibodies against it, detecting or

XX quantitating a polypeptide in tissue, as molecular weight markers and as

XX a food supplement (II) and its binding partners are useful in medical

XX imaging of sites expressing (II). (I) and (II) are useful for treating

XX disorders involving aberrant protein expression or biological activity.

XX The polypeptide and polynucleotide sequences have applications in

XX diagnostics, forensics, gene mapping, identification of mutations

CC responsible for genetic disorders or other traits to assess biodiversity

CC and to produce other types of data and products dependent on DNA and

CC amino acid sequences. AAS64197-AAS94564 represent novel human

CC diagnostic coding sequences of the invention.

CC Note: The sequence data for this patent did not appear in the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pcl_sequences.

XX Sequence 2489 BP; 573 A; 708 C; 704 G; 504 T; 0 other;

XX Alignment Scores:

Pred. No.:	0.525	Length:	2489
Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.09%	Indels:	0
DB:	23	Gaps:	0

US-09-397-967-16 (1-1099) x AAS85547 (1-2489)

OY 941 VALHISARGASPLEUALAALARGASNTLEUVAL 952

DB 1814 GTTCATCGGAGCTTGCGACCGAACAATCTGTGA 1849

RESULT 53

ABL91661
ID ABL91661 standard; DNA: 2784 BP.

XX ABL91661;

XX 28-MAY-2002 (first entry)

XX Human polynucleotide SEQ ID NO 4.

XX Human; HIV; HCV; gene expression; oligonucleotide; tumour; pathogen;

XX Plasmodium; virus; viroid; cytokine; prion; antisense oligonucleotide;

XX cyostatic; virucide; protozoacide; antibacterial; ds.

XX Homo sapiens.

XX DEL100586-C1.

XX 11-APR-2002.

XX 09-JAN-2001; 2001DE-1000586.

XX 09-JAN-2001; 2001DE-1000586.

XX (RIBO-) RIBOPHARMA AG.

XX Kreutzer R, Limmer S, Rost S, Hadwiger P;

XX WPI: 2002-270454/32.

XX Inhibiting gene expression in cells, useful for e.g. treating tumors,

XX by introducing double-stranded complementary oligoRNA having unpaired

XX terminal bases

XX Claim 13; Page 8-9; 104pp; German.

XX The invention relates to a method for inhibiting expression of a target

XX gene (ABL91658-ABL91797) in a cell by introducing at least one

XX oligonucleotide that has a double-stranded structure consisting of at

XX least 49 sequential nucleotide pairs, with at least part of one strand

XX complementary with the target gene and has at least one end a

XX single-stranded segment of 1-4 nt. The method provides

XX oligonucleotides for antisense inhibition of gene expression useful

XX e.g. for treating tumors but the oligonucleotides may also be

XX directed against genes present in pathogens (e.g. Plasmodium or

XX viruses/viroids, pathogenic on humans, animals or plants) or against

XX cytokine, Id, developmental or prion genes. The method provides more

XX effective inhibition of gene expression than use of known

XX oligonucleotides, probably because the unpaired overhang increases

CC stability and thus intracellular concentration.
 XX Sequence 2784 BP; 770 A; 650 C; 731 G; 633 T; 0 other;
 SQ

Alignment Scores:
 Pred. No.: 0.584 Length: 2784
 Score: 12.00 Matches: 12
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.09% Indels: 0
 DB: 24 Gaps: 0

US-09-397-967-16 (1-1099) x ABL91661 (1-2784)

OY 941 VALHISARGASPLEUALAALARGASNILEUVAL 952
 |||||||||||||||||||||||||||||||||||
 Db 2050 GTGCATCGTATCTGCGCGACGAGACATCTGTG 2085

RESULT 54
 AAQ90654
 ID AAQ90654 standard; cDNA; 2820 BP.
 XX
 AC AAQ90654;
 XX
 DT 11-NOV-1995 (first entry)
 XX
 DE Eph-related PTK Cdk8 CDNA.
 XX
 KM Cdk8; Eph; protein tyrosine-kinase; PTK; cancer; diagnosis;
 KM prognosis; ss.
 XX
 OS Gallus sp.
 XX

Key Location/Qualifiers
 FT CDS 2..2551
 FT /*tag= a

MO9515375-A.
 PN
 PD 08-JUN-1995.
 XX
 PF 07-SEP-1994; 94MO-US10140.
 XX
 PR 03-DEC-1993; 93US-0162809.
 XX
 PA (LJOL-) LA JOLLA CANCER RES FOUND.
 XX
 PI Pasquale EB, Sajjadi FG;
 DR WPI: 1995-215256/28.
 DR P-PSDB: AAR75706.
 XX

Eph-related protein tyrosine kinase(s) - for monitoring and diagnosing cancer.
 PT
 PS
 PS Disclosure: Page 50-53; 129pp; English.
 XX
 PS
 CC The chick Eph-related PTK cDNA clone given in AAQ90653 encoded Cdk8 (AAR75706) that showed 74% amino acid identity in the catalytic and C-terminal regions to Cdk5 (AAR75712) and novel Cdk6 (AAR75704), Cdk7 (AAR75705) and Cdk10 (AAR75708). Cdk8 was expressed in adult brain and retina and at lower levels in kidney, lung, skeletal muscle and thymus.
 CC
 CC
 XX
 SQ Sequence 2820 BP; 753 A; 693 C; 731 G; 643 T; 0 other;

Alignment Scores:
 Pred. No.: 0.591 Length: 2820
 Score: 12.00 Matches: 12
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.09% Indels: 0
 DB: 16 Gaps: 0

US-09-397-967-16 (1-1099) x AAQ90654 (1-2820)

OY 941 VALHISARGASPLEUALAALARGASNILEUVAL 952
 |||||||||||||||||||||||||||||||||||
 Db 1817 GTGCATCGGATCTAGCTCTCGAACAACATCTGTG 1852

RESULT 55
 ABL91658
 ID ABL91658 standard; DNA; 2955 BP.
 XX
 AC ABL91658;
 XX
 DT 28-MAY-2002 (first entry)
 XX
 DE Human polynucleotide SEQ ID NO 1.
 XX
 KM Human; HIV; HCV; gene expression; oligonucleotide; tumour; pathogen;
 KM Plasmidium; virus; viroid; cytokine; prion; antisense oligonucleotide;
 KM cytosolic; virucide; protozoacide; antibacterial; ds.
 XX
 OS Homo sapiens.
 XX
 PN DE10100586-C1.
 XX
 PD 11-APR-2002.
 XX
 PF 09-JAN-2001; 2001DE-1000586.
 XX
 PR 09-JAN-2001; 2001DE-1000586.
 XX
 PA (RIBO-) RIBOPHARMA AG.
 XX
 PI Kreutzer R, Lommer S, Rost S, Hadwiger P;
 DR WPI: 2002-270454/32.
 XX

Inhibiting gene expression in cells, useful for e.g. treating tumors, by introducing double-stranded complementary oligoRNA having unpaired terminal bases
 PT
 PT
 PS Claim 13; Page 5-6; 104pp; German.
 XX
 CC The invention relates to a method for inhibiting expression of a target gene (ABL91658-AB191797) in a cell by introducing at least one oligonucleotide that has a double-stranded structure consisting of at least 49 sequential nucleotide pairs, with at least part of one strand complementary with the target gene and has at least one end a single-stranded segment of 1-4 nt. The method provides oligonucleotides for antisense inhibition of gene expression useful e.g. for treating tumors but the oligonucleotides may also be directed against genes present in pathogens (e.g. plasmidium or viruses/viroids), pathogenic on humans, animals or plants) or against cytokine, id, developmental or prion genes. The method provides more effective inhibition of gene expression than use of known oligonucleotides, probably because the unpaired overhang increases stability and thus intracellular concentration.
 CC
 CC
 XX
 SQ Sequence 2955 BP; 597 A; 868 C; 904 G; 586 T; 0 other;

Alignment Scores:
 Pred. No.: 0.617 Length: 2955
 Score: 12.00 Matches: 12
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.09% Indels: 0
 DB: 24 Gaps: 0

US-09-397-967-16 (1-1099) x ABL91658 (1-2955)

OY 941 VALHISARGASPLEUALAALARGASNILEUVAL 952
 |||||||||||||||||||||||||||||||||||
 Db 2260 GTCCACGGGACCTGCTGCTCGAACAACATCTGTG 2295

KW Human; HIV; HCV; gene expression; oligoribonucleotide; tumour; pathogen;
KW Plasmodium; virus; viroid; cytokine; prion; antisense oligonucleotide;
KW Cytostatic; virucide; protozoacide; antibacterial; ds.
OS Homo sapiens.
XX DE10100586-Cl.
XX PD 11-APR-2002.
XX PF 09-JAN-2001; 2001DE-1000586.
XX PR 09-JAN-2001; 2001DE-1000586.
XX PA (RIBO-) RIBOPHARMA AG.
XX PI Kreutzer R, Limmer S, Rost S, Hadwiger P;
XX WPI: 2002-270454/32.
XX DR
XX PT Inhibiting gene expression in cells, useful for e.g. treating tumors,
XX by introducing double-stranded complementary oligoRNA having unpaired
XX terminal bases
XX PS
XX SQ Claim 13; Page 24-25; 104pp; German.
XX
XX The invention relates to a method for inhibiting expression of a target
XX gene (ABL91658-ABL91797) in a cell by introducing at least one
XX oligoribonucleotide that has a double-stranded structure consisting of at
XX most 49 sequential nucleotide pairs, with at least part of one strand
XX complementary with the target gene and has at least one end a
XX single-stranded segment of 1-4 nt. The method provides
XX oligoribonucleotides for antisense inhibition of gene expression useful
XX e.g. for treating tumors but the oligoribonucleotides may also be
XX directed against genes present in pathogens (e.g. Plasmodium or
XX viruses/viroids, pathogenic on humans, animals or plants) or against
XX cytokine, id. developmental or prion genes. The method provides more
XX effective inhibition of gene expression than use of known
XX oligonucleotides, probably because the unpaired overhang increases
XX stability and thus intracellular concentration.
XX
XX Alignment Scores:
XX Pred. No.: 0.619 Length: 2964
XX Score: 12.00 Matches: 12
XX Percent Similarity: 100.00% Conservative: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 1.09% Indels: 0
XX DB: 24 Gaps: 0
XX
XX US-09-397-967-16 (1-1099) x ABL91681 (1-2964)
XX
XX QY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
XX |||||||||||||||||||||||||||||||||||
XX DB 2209 GTCCACCGAGACTGGCTGCTGCGCAACATCTTAGTC 2244
XX
XX RESULT 59
XX ID AAO90971 standard; cDNA to mRNA; 2982 BP.
XX XX
XX AC AAO90971;
XX
XX DT 24-NOV-1995 (first entry)
XX
XX DE Protein p140 cDNA from rat skeletal muscle myoblast cell line L6.
XX
XX KW Protein p140: insulin; tyrosine phosphorylation; ss.
XX
XX OS Rattus rattus.
XX
XX PN FP659883-A.
XX
XX

PD 28-JUN-1995.
XX
XX PF 24-NOV-1994; 94EP-0118524.
XX
XX PR 24-NOV-1993; 93JP-0315806.
XX
XX PA (ONOV) ONO PHARM CO LTD.
XX
XX PI Kitagawa K, Ohno H, Tajima H;
XX
XX DR WPI: 1995-226291/30.
XX
XX PT Isolated protein p140 polypeptide - and treatment of diabetes based
XX on tyrosine phosphorylation of protein p140.
XX
XX PS Claim 4; Page 23-25; 42pp; English.
XX
XX CC p140 is used for the prevention and treatment of diabetes. Dosage
XX is 10 microg-1000 mg (p.o.) or 10 microg-100mg (i.v.). A
XX polypeptide of protein p140 having the sequence in AAR75843 is
XX claimed. also claimed is a DNA encoding it, specifically,
XX AAO90971 and AAO90972.
XX
XX SQ Sequence 2982 BP; 639 A; 889 C; 853 G; 601 T; 0 other;
XX
XX Alignment Scores:
XX Pred. No.: 0.623 Length: 2982
XX Score: 12.00 Matches: 12
XX Percent Similarity: 100.00% Conservative: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 1.09% Indels: 0
XX DB: 16 Gaps: 0
XX
XX US-09-397-967-16 (1-1099) x AAO90971 (1-2982)
XX
XX QY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
XX |||||||||||||||||||||||||||||||||||
XX DB 2248 GTGCACCGTAGCTGCTGCGCCGCAACATCTTAGTC 2283
XX
XX RESULT 60
XX ID ABL91662 standard; DNA; 2997 BP.
XX
XX AC ABL91662;
XX
XX DT 28-MAY-2002 (first entry)
XX
XX DE Human polynucleotide SEQ ID NO. 5.
XX
XX KW Human; HIV; HCV; gene expression; oligoribonucleotide; tumour; pathogen;
KW Plasmodium; virus; viroid; cytokine; prion; antisense oligonucleotide;
KW Cytostatic; virucide; protozoacide; antibacterial; ds.
XX
XX OS Homo sapiens.
XX
XX PN DE10100586-Cl.
XX
XX PD 11-APR-2002.
XX
XX PF 09-JAN-2001; 2001DE-1000586.
XX
XX PR 09-JAN-2001; 2001DE-1000586.
XX
XX PA (RIBO-) RIBOPHARMA AG.
XX
XX PI Kreutzer R, Limmer S, Rost S, Hadwiger P;
XX
XX DR WPI: 2002-270454/32.
XX
XX PT Inhibiting gene expression in cells, useful for e.g. treating tumors,
XX by introducing double-stranded complementary oligoRNA having unpaired
XX terminal bases
XX

PS Claim 13: Page 9-10; 104pp; German.
XX
CC The invention relates to a method for inhibiting expression of a target
CC gene (ABL91658-ABL91797) in a cell by introducing at least one
CC oligoribonucleotide that has a double-stranded structure consisting of at
CC most 49 sequential nucleotide pairs, with at least part of one strand
CC complementary with the target gene and has at least one end a
CC single-stranded segment of 1-4 nt. The method provides
CC oligoribonucleotides for antisense inhibition of gene expression useful
CC e.g. for treating tumours but the oligoribonucleotides may also be
CC directed against genes present in pathogens (e.g. Plasmodium or
CC viruses/viroids, pathogenic on humans, animals or plants) or against
CC cytokine, id, developmental or prion genes. The method provides more
CC effective inhibition of gene expression than use of known
CC oligonucleotides, probably because the unpaired overhang increases
CC stability and thus intracellular concentration.
SQ Sequence 2997 BP; 901 A; 593 C; 732 G; 771 T; 0 other:

Alignment Scores:
Pred. No.: 0.625 Length: 2997
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 24 Gaps: 0

US-09-397-967-16 (1-1099) x ABL91662 (1-2997)
OY 941 VALHISARGSPLEUALAALARGASNIIELEUVAL 952
DB 2263 GTTCACGCGGACCTGCTGCACATATCTTGTTC 2298
RESULT 61
ABL91680
ID ABL91680 standard; DNA: 2997 BP.
XX
AC ABL91680;
XX
DT 28-MAY-2002 (first entry)
XX
DE Human polynucleotide SEQ ID NO 23.
XX
XX
KW Human; HIV; HCV; gene expression; oligoribonucleotide; tumour; pathogen;
KW Plasmodium; virus; viroid; cytokine; prion; antisense oligonucleotide;
KW cytostatic; virucide; protozoacide; antibacterial; ds.
XX
OS Homo sapiens.
XX
PN DE10100586-C1.
XX
PD 11-APR-2002.
XX
PE 09-JAN-2001; 2001DE-1000586.
XX
PR 09-JAN-2001; 2001DE-1000586.
XX
PA (RIBO-) RIBOPHARMA AG.
XX
PI Kreutzer R, Limmer S, Rost S, Hadwiger P;
XX
DR WPI: 2002-270454/32.
XX
PT Inhibiting gene expression in cells, useful for e.g. treating tumours,
PT by introducing double-stranded complementary oligoRNA having unpaired
XX terminal bases
XX
PS Claim 13: Page 23-24; 104pp; German.
XX
CC The invention relates to a method for inhibiting expression of a target
CC gene (ABL91658-ABL91797) in a cell by introducing at least one
CC oligoribonucleotide that has a double-stranded structure consisting of at
CC most 49 sequential nucleotide pairs, with at least part of one strand

CC complementary with the target gene and has at least one end a
CC single-stranded segment of 1-4 nt. The method provides
CC oligoribonucleotides for antisense inhibition of gene expression useful
CC e.g. for treating tumours but the oligoribonucleotides may also be
CC directed against genes present in pathogens (e.g. Plasmodium or
CC viruses/viroids, pathogenic on humans, animals or plants) or against
CC cytokine, id, developmental or prion genes. The method provides more
CC effective inhibition of gene expression than use of known
CC oligonucleotides, probably because the unpaired overhang increases
CC stability and thus intracellular concentration.
SQ Sequence 2997 BP; 590 A; 940 C; 875 G; 592 T; 0 other:

Alignment Scores:
Pred. No.: 0.625 Length: 2997
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 24 Gaps: 0

US-09-397-967-16 (1-1099) x ABL91680 (1-2997)
OY 941 VALHISARGSPLEUALAALARGASNIIELEUVAL 952
DB 2263 GTTCACGCGGACCTGCTGCACATCTTGTTC 2298
RESULT 62
ABL91659
ID ABL91659 standard; DNA: 3042 BP.
XX
AC ABL91659;
XX
DT 28-MAY-2002 (first entry)
XX
DE Human polynucleotide SEQ ID NO 2.
XX
XX
KW Human; HIV; HCV; gene expression; oligoribonucleotide; tumour; pathogen;
KW Plasmodium; virus; viroid; cytokine; prion; antisense oligonucleotide;
KW cytostatic; virucide; protozoacide; antibacterial; ds.
XX
OS Homo sapiens.
XX
PN DE10100586-C1.
XX
PD 11-APR-2002.
XX
PE 09-JAN-2001; 2001DE-1000586.
XX
PR 09-JAN-2001; 2001DE-1000586.
XX
PA (RIBO-) RIBOPHARMA AG.
XX
PI Kreutzer R, Limmer S, Rost S, Hadwiger P;
XX
DR WPI: 2002-270454/32.
XX
PT Inhibiting gene expression in cells, useful for e.g. treating tumours,
PT by introducing double-stranded complementary oligoRNA having unpaired
XX terminal bases
XX
PS Claim 13: Page 6-7; 104pp; German.
XX
CC The invention relates to a method for inhibiting expression of a target
CC gene (ABL91658-ABL91797) in a cell by introducing at least one
CC oligoribonucleotide that has a double-stranded structure consisting of at
CC most 49 sequential nucleotide pairs, with at least part of one strand
CC complementary with the target gene and has at least one end a
CC single-stranded segment of 1-4 nt. The method provides
CC oligoribonucleotides for antisense inhibition of gene expression useful
CC e.g. for treating tumours but the oligoribonucleotides may also be
CC directed against genes present in pathogens (e.g. Plasmodium or
CC viruses/viroids, pathogenic on humans, animals or plants) or against

CC cytokine, Id, developmental or prion genes. The method provides more
 CC effective inhibition of gene expression than use of known
 CC oligonucleotides, probably because the unpaired overhang increases
 CC stability and thus intracellular concentration.
 XX

SO Sequence 3042 BP; 600 A; 941 C; 960 G; 541 T; 0 other;

Alignment Scores:

Pred. No.:	0.634	Length:	3042
Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.09%	Indels:	0
DB:	24	Gaps:	0

US-09-397-967-16 (1-1099) x ABL91659 (1-3042)

Oy 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
 |||||
 Db 2317 GTGCACCGTGCCTGCTGCCGCAACATCCTCCTC 2352

RESULT 63

AA07308	Location/Qualifiers
ID AAT07308 standard; CDNA; 3105 BP.	
XX AC AAT07308;	
DT 19-MAR-1996 (first entry)	
XX DE Receptor tyrosine kinase (neural kinase) cDNA.	
XX KW Receptor tyrosine kinase; neural kinase; Nuk gene; axon;	
XX KW axonogenesis; nerve disorder; gene therapy; transgenic animal; ss.	
XX OS Mus musculus.	
XX FH Key	Location/Qualifiers
FT CDS	1..2985
FT	/*tag= a
FT sig_peptide	1..78
FT	/*tag= b
FT mat_peptide	79..2982
FT	/*tag= c
XX PN W09530326-A1.	
XX PD 09-NOV-1995.	
XX PF 28-APR-1995; 95WO-CA00254.	
XX PR 29-APR-1994; 94US-0235407.	
XX PA (MOUN) MOUNT SINAI HOSPITAL CORP.	
XX PI Henkemeyer M, Letwin K, Pawson A;	
XX DR WPI; 1995-393299/50.	
XX DR P-PSDB; AAR87018.	
XX PT DNA encoding neural receptor tyrosine kinase - useful in gene	
XX PT therapy of nerve disorders, and for diagnosis and identification of	
XX PT therapeutic agents	
XX PS Claim 2; Page 68-70; 103pp; English.	
XX CC cDNA clones pNukRACE A2 and K2 were combined to obtain a sequence	
XX CC (AA07308) coding for a novel receptor tyrosine kinase, designated	
XX CC neural kinase (Nuk) (AAR87018). The clones were obtd. from a cDNA	
XX CC library of mouse embryo cDNA in lambda-gli10 probed with a partial Nuk	
XX CC cDNA insert. The gene was mapped to the distal end of chromosome 4	
XX CC near the and-1 mutation. The cDNA is used to produce recombinant	
XX CC Nuk or transgenic animal models for studies of Nuk function, or as	
XX CC probes to detect genes or diagnose conditions associated with Nuk	

CC gene mutation, or in gene (antisense) therapy of disorders of the
 CC nervous system.
 XX

SO Sequence 3105 BP; 711 A; 937 C; 847 G; 610 T; 0 other;

Alignment Scores:

Pred. No.:	0.647	Length:	3105
Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.09%	Indels:	0
DB:	16	Gaps:	0

US-09-397-967-16 (1-1099) x AAT07308 (1-3105)

Oy 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
 |||||
 Db 2251 GTGCACCGTGCCTGCTGCYGAACATCCTCCTC 2286

RESULT 64

AA084528	Location/Qualifiers
ID AAT84528 standard; CDNA; 3105 BP.	
XX AC AAT84528;	
DT 02-DEC-1997 (first entry)	
XX DE Mouse Nuk tyrosine kinase cDNA.	
XX KW Nuk tyrosine kinase; Eph receptor tyrosine kinase;	
XX KW signal transduction; axonogenesis; neurodegenerative disease;	
XX KW Alzheimer's disease; Parkinson's disease; Huntington's disease;	
XX KW multiple sclerosis; amyotrophic lateral sclerosis;	
XX KW Wernicke's disease; nerve damage; trauma; ischemia; stroke; ss.	
XX OS Mus musculus.	
XX FH Key	Location/Qualifiers
FT CDS	1..2985
FT	/*tag= a
FT sig_peptide	1..78
FT	/*tag= b
FT mat_peptide	79..2982
FT	/*tag= c
XX PN W09714966-A1.	
XX PD 24-APR-1997.	
XX PF 10-OCT-1996; 96WO-CA00679.	
XX PR 13-OCT-1995; 95US-0005518.	
XX PA (MOUN) MOUNT SINAI HOSPITAL CORP.	
XX PI Henkemeyer M, Pawson A;	
XX DR WPI; 1997-245245/22.	
XX DR P-PSDB; AAW26366.	
XX PT Activation of ligand regulatory pathways by Eph subfamily receptor	
XX PT tyrosine kinases - for stimulating or inhibiting axonogenesis,	
XX PT useful for treatment of e.g. neurodegenerative diseases such as	
XX PT Alzheimer's or Parkinson's diseases	
XX PS Disclosure; Page 24-25; 55pp; English.	
XX CC This cDNA sequence codes for murine Nuk tyrosine kinase (AAW26366),	
XX CC an Eph subfamily receptor tyrosine kinase essential for formation	
XX CC of the medial tract of the anterior commissure of the brain. Its	
XX CC sequence was deduced from clones isolated from an embryo cDNA	
XX CC library in lambda gli10. The Nuk gene maps to the distal end of	
XX CC chromosome 4 near the and-1 mutation. Eph subfamily receptor	

CC tyrosine kinases (e.g. Nuk extracellular domain polypeptides)
 CC can be used in claimed methods to: activate a ligand regulatory
 CC pathway in a cell; identify a substance able to bind a ligand for
 CC an Eph subfamily receptor tyrosine kinase; and to affect neuronal
 CC development or regeneration, especially the stimulation or
 CC inhibition of axonogenesis, in a mammal. Activation of the ligand
 CC regulatory pathways results in downstream activation of a series of
 CC division, cytoskeletal architecture, cell metabolism, cell
 CC migration and cell-cell interactions. Substances which activate
 CC the ligand regulatory pathway may be used for stimulating or
 CC inhibiting neuronal development, regeneration and axonal migration
 CC associated with neurodegenerative disease e.g. Alzheimer's,
 CC Parkinson's or Huntington's diseases, multiple sclerosis,
 CC amyotrophic lateral sclerosis, deficiency diseases such as
 CC Wernicke's disease, peripheral nerve damage, trauma and ischaemia
 CC resulting from stroke.

SQ Sequence 3105 BP: 710 A; 937 C; 848 G; 610 T; 0 other;

Alignment Scores:

Prod. No.:	0.647	Length:	3105
Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.09%	Indels:	0
DB:	18	Gaps:	0

US-09-397-967-16 (1-1099) x AAT84538 (1-3105)

OY 941 VALHISARGASPLEUALAIAARGASNIJLeuVal 952

DB 2251 GTGCACCGTGACCTTGCTGCGAACAACATCCTCCTC 2286

RESULT 65

AAS03812

ID AAS03812 standard; cDNA: 3105 BP.

XX AAS03812;

XX 29-AUG-2001 (first entry)

DE Murine neural kinase (Nuk) cDNA.

KW Neural kinase; Nuk; receptor tyrosine kinase; axonal migration; stroke;
 KW nerve fibre; cell-cell interaction; axonogenesis; neuronal development;
 KW regeneration; neurodegenerative disorder; Alzheimer's disease; ischaemia;
 KW Parkinson's disease; Huntington's disease; demyelinating disease; ss;
 KW multiple sclerosis; amyotrophic lateral sclerosis; deficiency disease;
 KW Wernicke's disease; nutritional polyneuropathy; multistem degeneration;
 KW progressive supranuclear palsy; Shy Drager's syndrome; mouse;
 KW olivoponto cerebellar atrophy; peripheral nerve damage.

XX Mus musculus.

XX Key Location/Qualifiers

XX CDS

XX 1..2985

XX /tag- a

XX /product- "Murine neural kinase"

XX sig-peptide

XX 1..78

XX /tag- b

XX mat-peptide

XX 79..2982

XX /tag- c

XX /product- "Murine neural kinase"

XX US6218356-B1.

XX 17-APR-2001.

XX 13-OCT-1995.

XX 95US-0542635.

XX 29-APR-1994.

XX 94US-0235407.

XX 28-APR-1995.

XX (MOUN) MOUNT SINAI HOSPITAL CORP.
 PA Pawsen A, Henkemeyer M, Letwin K;
 XX WPI: 2001-289845/30.
 DR P-PSDB: AAU01907.
 XX
 PT New composition comprising neural receptor tyrosine kinase protein
 PT useful for inhibiting or stimulating axonogenesis, neuronal
 PT development, or regeneration and axonal migration
 PS Disclosure: Fig 1; 86pp; English.

The sequence represents a cDNA molecule which encodes a mouse neural
 CC kinase (Nuk) polypeptide. The murine Nuk locus has been mapped to the
 CC distal end of chromosome four. The polypeptide is a novel receptor
 CC tyrosine kinase protein, and is found to be expressed at high levels
 CC within migrating axons and is associated with nerve fibres. It functions
 CC to regulate specific cell-cell interactions during early development of
 CC the nervous system and in axonogenesis. Substances which bind to the Nuk
 CC protein, particularly ligands, may be used for stimulating or inhibiting
 CC neuronal development, regeneration and axonal migration associated with
 CC neurodegenerative disorders and conditions involving trauma and injury to
 CC the nervous system. These disorders include Alzheimer's disease,
 CC Parkinson's disease, Huntington's disease, demyelinating diseases such as
 CC multiple sclerosis, amyotrophic lateral sclerosis, deficiency diseases
 CC such as Wernicke's disease and nutritional polyneuropathy, progressive
 CC ponto cerebellar atrophy, Shy Drager's syndrome, multistem degeneration, olivop
 CC resulting from stroke. The proteins may be used to prepare antibodies
 CC having specificity for Nuk proteins, which can be used to diagnose or
 CC treat disorders of the nervous system. These proteins are also used for
 CC screening agonists or antagonists of the interactions of the Nuk proteins
 CC with binding molecules.

SQ Sequence 3105 BP: 710 A; 937 C; 848 G; 610 T; 0 other;

Alignment Scores:

Pred. No.:	0.647	Length:	3105
Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.09%	Indels:	0
DB:	22	Gaps:	0

US-09-397-967-16 (1-1099) x AAS03812 (1-3105)

OY 941 VALHISARGASPLEUALAIAARGASNIJLeuVal 952

DB 2251 GTGCACCGTGACCTTGCTGCGAACAACATCCTCCTC 2286

RESULT 66

AAS69876

ID AAS69876 standard; cDNA: 3107 BP.

XX AAS69876;

XX 13-FEB-2002 (first entry)

DE DNA encoding novel human diagnostic protein #5680.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US08631.

13-DEC-2001.
30-MAY-2001; 2001WO-US10838.
05-JUN-2000; 2000US-209473P.
05-JUN-2000; 2000US-209531P.
18-SEP-2000; 2000US-233133P.
18-SEP-2000; 2000US-233617P.
20-SEP-2000; 2000US-234009P.
20-SEP-2000; 2000US-234034P.
20-SEP-2000; 2000US-234052P.
20-SEP-2000; 2000US-234509P.
22-SEP-2000; 2000US-234567P.
22-SEP-2000; 2000US-234924P.
25-SEP-2000; 2000US-235072P.
25-SEP-2000; 2000US-235082P.
25-SEP-2000; 2000US-235134P.
25-SEP-2000; 2000US-235280P.
26-SEP-2000; 2000US-235638P.
26-SEP-2000; 2000US-235711P.
27-SEP-2000; 2000US-235720P.
27-SEP-2000; 2000US-235840P.
27-SEP-2000; 2000US-235863P.
28-SEP-2000; 2000US-236028P.
28-SEP-2000; 2000US-236033P.
28-SEP-2000; 2000US-236034P.
28-SEP-2000; 2000US-236109P.
28-SEP-2000; 2000US-236111P.
29-SEP-2000; 2000US-236891P.
29-SEP-2000; 2000US-237172P.
02-OCT-2000; 2000US-237173P.
02-OCT-2000; 2000US-237278P.
02-OCT-2000; 2000US-237294P.
02-OCT-2000; 2000US-237295P.
02-OCT-2000; 2000US-237315P.
03-OCT-2000; 2000US-237425P.
03-OCT-2000; 2000US-237598P.
03-OCT-2000; 2000US-237604P.
03-OCT-2000; 2000US-237606P.
03-OCT-2000; 2000US-237608P.
01-NOV-2000; 2000US-244857P.
01-NOV-2000; 2000US-245084P.
(AVAL-) AVALON PHARM.
Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
Soppet DR, Weaver Z;
WPI: 2002-188264/24.
Screening for anti-neoplastic agent involves exposing cells to a
chemical agent to be tested for anti-neoplastic activity, and
determining a change in expression of a gene of a signature gene set
Claim 1; SEQ ID 4155; 44pp; English.
The present invention describes a method (M1) for screening for an
anti-neoplastic agent. The method involves exposing cells to a chemical
agent to be tested for anti-neoplastic activity, determining a change in
expression of at least one gene (1) of a signature gene set, where (1)
comprises a sequence (S) selected from 8447 sequences (given in ABL6166
to ABL70110), or is at least 95% identical to (S), where a change in
expression is indicative of anti-neoplastic activity. (1) has cytostatic
activity and can be used in gene therapy. M1 can be used for screening
an anti-neoplastic agent, and can be used for producing a product which
is the data collected with respect to the anti-neoplastic agent as a
result of M1, and the data is sufficient to convey the chemical
structure and/or properties of the agent. M1 can be used in the

CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
 CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
 CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
 CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
 CC carcinoma, papillary carcinoma and Wilms' tumour.
 XX

SO Sequence 3107 BP; 850 A; 720 C; 812 G; 725 T; 0 other;

Alignment Scores:

Pred. No.:	Length:	Matches:
Score: 0.647	3107	12
Percent Similarity: 12.00		
Best Local Similarity: 100.00%		
Query Match: 100.00%		
DB: 1.09%		
Indels: 0		
Gaps: 0		

US-09-397-967-16 (1-1099) x ABL65818 (1-3107)

OY 941 ValHisArgAspLeuAlaAlaArgAsnIleuVal 952

DB 2260 GTGCATCGTCGATCTGCGCCGACGAGACATCTCTGGTG 2295

RESULT 68

ID AAT02948 standard; cDNA; 3116 BP.

XX AAT02948;

DT 16-APR-1996 (first entry)

DE Eph-like receptor protein tyrosine kinase HEK8 cDNA.

KW Eph-like receptor protein tyrosine kinase; PTK; HEK8;

KM human eph-like kinase; therapy; diagnosis; antibody; vector; ss.

OS Homo sapiens.

XX Key Location/Qualifiers

FT CDS 34..2994

XX /*tag= a

PN MO9528484-A1.

XX 26-OCT-1995.

PF 14-APR-1995; 95WO-US04681.

XX 15-APR-1994; 94US-0229509.

XX (AMGE-) AMGEN INC.

PI Fox GM, Jing S, Welcher AA;

DR WPI; 1995-373799/48.

DR P-PSDB; AAR85091.

XX New nucleic acid encoding Eph-like receptor tyrosine kinase(s)
 PT and related vectors, host cells, proteins, antibodies etc., used
 PT diagnostically and therapeutically to modulate receptor activation
 or prodn.

XX Claim 1; Page 57-62; 133pp; English.

CC cDNAs (AAT02946-49) coding for 4 novel human Eph-like receptor protein
 CC tyrosine kinases, HEK5, HEK7, HEK8 and HEK11 (AAR85089-92).

CC respectively, were isolated from a human foetal brain cDNA library using
 CC a directed PCR approach with primers (see AAT02960-61) based on conserved
 CC regions of receptor PTKs and Eph-like receptor PTKs. HEK5, HEK7 and HEK8
 CC show extensive homology to the catalytic domain of chicken Eph-like
 CC receptors Cek5, Cek7 and Cek8. HEK11 shows no homology to any known
 CC Eph-like receptor. The isolated cDNAs are used for prodn. of
 CC recombinant HEKs and chimeric receptors, in hybridisation assays, and
 CC to detect abnormalities in HEK receptor genes.

XX
 SO Sequence 3116 BP; 859 A; 720 C; 812 G; 725 T; 0 other;

Alignment Scores:

Pred. No.:	Length:	Matches:
Score: 0.649	3116	12
Percent Similarity: 12.00		
Best Local Similarity: 100.00%		
Query Match: 100.00%		
DB: 1.09%		
Indels: 0		
Gaps: 0		

US-09-397-967-16 (1-1099) x AAT02948 (1-3116)

OY 941 ValHisArgAspLeuAlaAlaArgAsnIleuVal 952

DB 2260 GTGCATCGTCGATCTGCGCCGACGAGACATCTCTGGTG 2295

RESULT 69

ID AAG90652 standard; cDNA; 3133 BP.

XX AAG90652;

DT 11-NOV-1995 (first entry)

DE Eph-related tyrosine kinase CEK6 cDNA.

KW Cek6; Eph: protein tyrosine-kinase; PTK; cancer; diagnosis;

KM prognosis; ss.

OS Gallus sp.

XX Key Location/Qualifiers

FT CDS 3..419

FT CDS /*tag= a

XX 421..2859

XX /*tag= b

PN MO9515375-A.

XX 08-JUN-1995.

PF 07-SEP-1994; 94WO-US10140.

XX 03-DEC-1993; 93US-0162809.

XX (LJOL-) LA JOLLA CANCER RES FOUND.

PI Pasquale EB, Sajjadi FG;

DR WPI; 1995-215256/28.

DR P-PSDB; AAR75704.

XX Eph-related protein tyrosine kinase(s) - for monitoring and diagnosing
 PT cancer.
 PT Disclosure: Page 37-41; 129pp; English.

CC Novel Eph-related PTK Cek6 cDNA clones (AAG90652) were isolated from
 CC chick embryo and embryonic brain cDNA libraries in phage lambda gt11.
 CC The encoded Cek6 protein (AAR75704) is closely related to rat Elk,
 CC Ceks (AAR75712) and Cek10 (AAR75708). Cek6 transcripts were found in
 CC 10-day embryos and in adult brain, lung, heart and skeletal muscle.

SO Sequence 3133 BP; 718 A; 918 C; 922 G; 575 T; 0 other;

Alignment Scores:

Pred. No.:	Length:	Matches:
Score: 0.652	3133	12
Percent Similarity: 12.00		
Best Local Similarity: 100.00%		
Query Match: 100.00%		
DB: 1.09%		
Indels: 0		
Gaps: 0		

US-09-397-967-16 (1-1099) x AA090652 (1-3133)

OY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
 DB 2125 GTCACAGGATCTGCGCCAGAACATCTCTGTC 2160

RESULT 70

AAA09322 standard; DNA: 3151 BP.

ID AAA09322:

AC AAA09322:

DT 10-AUG-2000 (first entry)

DE Human cancer associated antigen precursor DNA, clone NY-REN-47.

KM renal cancer; cancer associated antigen precursor; diagnosis;

KW cytosolic; ERK tyrosine kinase; ss.

OS Homo sapiens.

PN WO200020587-A2.

PD 13-APR-2000.

PF 04-OCT-1999; 99WO-US22873.

PR 05-OCT-1998; 98US-0166300.

PR 05-OCT-1998; 98US-0166350.

XX (LUDW-) LUDWIG INST CANCER RES.

PI Obata Y, Gout I, Tureci O, Sahin U, Pfeundschuh M, Scanlan MJ;

PI Stockert E, Chen Y, Old LJ, Jager E, Knuth A;

PT abnormal expression of human cancer associated antigens

PS Claim 57, Page 93-94; 121pp: English.

XX AAA09321-45 were isolated by SEREX screening from a renal cancer

CC cell line 1973/10.4. Homology searching revealed that these clones

CC ERK tyrosine kinase gene. The genes encode cancer associated antigen

CC precursors. These gene products are useful in methods for preventing

CC abnormal expression of human cancer associated antigens. The method

CC comprises contacting a sample from a subject with an agent that

CC specifically binds to the nucleic acid molecule or expression product

CC (or fragment) complexed with a human leukocyte antigen (HLA) molecule

CC and determining the interaction between the agent and the nucleic acid

CC molecule or the expression product as a determination of the disorder.

XX Sequence 3151 BP; 707 A; 974 C; 874 G; 596 T; 0 other;

Alignment Scores: 0.656 Length: 3151
 Pred. No.: 12.00 Matches: 12
 Score: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.09% Indels: 0
 DB: 21 Gaps: 0

US-09-397-967-16 (1-1099) x AAA09322 (1-3151)

OY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952

DB 2236 GTTCACCGTACCTGCTGCCGCAACATCTCTGTC 2271

RESULT 71

ABL91679 standard; DNA: 3168 BP.

AC ABL91679:

DT 28-MAY-2002 (first entry)

DE Human polynucleotide SEQ ID NO 22.

KM Human; HIV; HCV; gene expression; oligonucleotide; tumour; pathogen;

KW Plasmidium; virus; viroid; cytokine; prion; antisense oligonucleotide;

KW cytosolic; virucide; protozoacide; antibacterial; ds.

OS Homo sapiens.

PN DE10100586-C1.

PD 11-APR-2002.

PF 09-JAN-2001; 2001DE-1000586.

PR 09-JAN-2001; 2001DE-1000586.

PA (RIBO-) RIBOPHARMA AG.

PI Kreutzler R, Limmer S, Rost S, Hadwiger P;

PI WPI; 2002-270454/32.

PT Inhibiting gene expression in cells, useful for e.g. treating tumors,

PT by introducing double-stranded complementary oligoRNA having unpaired

PS Claim 13, Page 22-23; 104pp: German.

XX The invention relates to a method for inhibiting expression of a target

CC gene (ABL91658-ABL91797) in a cell by introducing at least one

CC oligonucleotide that has a double-stranded structure consisting of at

CC most 49 sequential nucleotide pairs, with at least part of one strand

CC complementary with the target gene and has at least one end a

CC single-stranded segment of 1-4 nt. The method provides

CC oligonucleotides for antisense inhibition of gene expression useful

CC e.g. for treating tumors but the oligonucleotides may also be

CC directed against genes present in pathogens (e.g. Plasmodium or

CC cytochrome, Id. development of gene expression than use of known

CC oligonucleotides, probably because the unpaired overhang increases

CC stability and thus intracellular concentration.

XX Sequence 3168 BP; 732 A; 947 C; 902 G; 587 T; 0 other;

Alignment Scores: 0.659 Length: 3168
 Pred. No.: 12.00 Matches: 12
 Score: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.09% Indels: 0
 DB: 24 Gaps: 0

US-09-397-967-16 (1-1099) x ABL91679 (1-3168)

OY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952

DB 2227 GTTCACCGTACCTGCTGCCGCAACATCTCTGTC 2262

RESULT 72

AAS87118 standard; CDNA: 3342 BP.

ID AAS87118:

DT 13-FEB-2002 (first entry)

XX DNA encoding novel human diagnostic protein #22922.
 DE Human: chromosome mapping; gene mapping; gene therapy; forensic;
 XX food supplement; medical imaging; diagnostic; genetic disorder; ss.
 KW Homo sapiens.
 XX MO200175067-A2.
 XX 11-OCT-2001.
 PD 30-MAR-2001; 2001WO-US08631.
 PF 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX (HWE-) HWEQ INC.
 PA Drmanac RT, Liu C, Tang YT;
 PI WPI: 2001-639362/73.
 DR P-PSDB: A8622931.
 XX New isolated polynucleotide and encoded polypeptides; useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX
 PS Claim 1: SEQ ID No 22922: 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human
 CC diagnostic coding sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 3342 BP; 898 A; 796 C; 867 G; 781 T; 0 other;
 XX
 Alignment Scores:
 Pred. No.: 0.693 Length: 3342
 Score: 12.00 Matches: 12
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.09% Indels: 0
 DB: 23 Gaps: 0
 US-09-397-967-16 (1-1099) x AAS87118 (1-3342)
 QY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
 DB 2494 GTGCATCGTATCTGGCCGACGAGACATCTGTG 2529
 RESULT 73
 AAT03100
 ID AAT03100 standard; DNA: 3348 BP.
 XX

AC AAT03100;
 XX 14-FEB-1996 (first entry).
 DE Protein tyrosine-kinase bptk7 gene.
 XX Protein tyrosine-kinase: ptk; bptk7; agonist; cell growth;
 KW differentiation; ss.
 XX Homo sapiens.
 OS
 XX
 FH Key Location/Qualifiers
 FT CDS 1..2961
 FT sig_peptide /*tag= a
 FT mat_peptide /*tag= b
 FT /*tag= c
 XX
 PN WO9527061-A1.
 PD 12-OCT-1995.
 PF 04-APR-1995; 95WO-US04228.
 XX
 PR 04-APR-1994; 94US-0222616.
 XX
 PA (GETH) GENENTECH INC.
 XX
 PI Bennett BD, Goeddel D, Lee JM, Matthews W, Tsai SP;
 PI Wood WI;
 DR WPI: 1995-366160/47.
 DR P-PSDB: AAR85936.
 XX
 PT Agonist antibodies which activate specific protein tyrosine
 PT kinase(s) - also activate chimeric proteins of kinase extracellular
 PT domain and Ig constant domain, useful for studying, and therapeutic
 PT modulation of, cell growth and differentiation
 PS
 XX Disclosure: Page 88-92; 125pp; English.
 XX
 CC DNA probes based on protein tyrosine-kinase (ptk) sequences were used
 CC to screen cDNA libraries to identify novel ptk genes. The bptks,
 CC bptk1, bptk2, bptk3, bptk4, bptk5 and bptk7 (AAR85924-28 and AAR85935,
 CC respectively) are expressed in human brain tissue and show homology
 CC to known ptk's. A full-length sequence for the bptk7 gene (AAT03100)
 CC was obt'd. This gene may be used to design new drugs, peptides and
 CC antisense constructs that modulate ptk activity.
 XX
 SQ Sequence 3348 BP; 924 A; 769 C; 855 G; 800 T; 0 other;
 XX
 Alignment Scores:
 Pred. No.: 0.694 Length: 3348
 Score: 12.00 Matches: 12
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.09% Indels: 0
 DB: 16 Gaps: 0
 US-09-397-967-16 (1-1099) x AAT03100 (1-3348)
 QY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
 DB 2227 GTGCATCGTATCTGGCCGACGAGACATCTGTG 2262
 RESULT 74
 ABL67952
 ID ABL67952 standard; DNA: 3370 BP.
 XX
 AC ABL67952;
 XX
 DT 15-MAY-2002 (first entry)
 XX

XX Ovary cancer related gene sequence SEQ ID NO:6289.
 DE
 XX Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
 KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
 KW cytosolic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
 XX gene; ds.
 OS Homo sapiens.
 XX
 PN WO200194629-A2.
 PD
 XX 13-DEC-2001.
 XX
 PF 30-MAY-2001; 2001WO-US10838.
 XX
 PR 05-JUN-2000; 2000US-209473P.
 PR 05-JUN-2000; 2000US-209533P.
 PR 18-SEP-2000; 2000US-233133P.
 PR 18-SEP-2000; 2000US-233617P.
 PR 20-SEP-2000; 2000US-234009P.
 PR 20-SEP-2000; 2000US-234034P.
 PR 20-SEP-2000; 2000US-234052P.
 PR 22-SEP-2000; 2000US-234509P.
 PR 22-SEP-2000; 2000US-234567P.
 PR 25-SEP-2000; 2000US-234923P.
 PR 25-SEP-2000; 2000US-234924P.
 PR 25-SEP-2000; 2000US-235077P.
 PR 25-SEP-2000; 2000US-235082P.
 PR 25-SEP-2000; 2000US-235134P.
 PR 25-SEP-2000; 2000US-235280P.
 PR 26-SEP-2000; 2000US-235637P.
 PR 26-SEP-2000; 2000US-235638P.
 PR 27-SEP-2000; 2000US-235711P.
 PR 27-SEP-2000; 2000US-235720P.
 PR 27-SEP-2000; 2000US-235840P.
 PR 27-SEP-2000; 2000US-235863P.
 PR 28-SEP-2000; 2000US-236028P.
 PR 28-SEP-2000; 2000US-236032P.
 PR 28-SEP-2000; 2000US-236033P.
 PR 28-SEP-2000; 2000US-236034P.
 PR 28-SEP-2000; 2000US-236109P.
 PR 28-SEP-2000; 2000US-236111P.
 PR 29-SEP-2000; 2000US-236842P.
 PR 29-SEP-2000; 2000US-236891P.
 PR 02-OCT-2000; 2000US-237172P.
 PR 02-OCT-2000; 2000US-237173P.
 PR 02-OCT-2000; 2000US-237278P.
 PR 02-OCT-2000; 2000US-237294P.
 PR 02-OCT-2000; 2000US-237295P.
 PR 02-OCT-2000; 2000US-237316P.
 PR 03-OCT-2000; 2000US-237425P.
 PR 03-OCT-2000; 2000US-237598P.
 PR 03-OCT-2000; 2000US-237604P.
 PR 03-OCT-2000; 2000US-237606P.
 PR 03-OCT-2000; 2000US-237608P.
 PR 01-NOV-2000; 2000US-244867P.
 PR 01-NOV-2000; 2000US-245084P.
 XX
 PA (AVALON PHARM.
 XX
 PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
 PI Soppet DR, Weaver Z;
 XX
 DR WPI; 2002-188264/24.
 XX
 PT Screening for anti-neoplastic agent involves exposing cells to a
 PT chemical agent to be tested for anti-neoplastic activity, and
 PT determining a change in expression of a gene of a signature gene set
 XX
 XX Claim 1; SEQ ID 6289; 44pp; English.
 PS
 CC The present invention describes a method (M1) for screening for an

CC anti-neoplastic agent. The method involves exposing cells to a chemical
 CC agent to be tested for anti-neoplastic activity, determining a change in
 CC expression of at least one gene (I) of a signature gene set, where (I)
 CC comprises a sequence (S) selected from 8447 sequences (given in AB61664
 CC to ABL70110), or is at least 95% identical to (S), where a change in
 CC expression is indicative of anti-neoplastic activity. (II) has cytosolic
 CC activity and can be used in gene therapy. M1 can be used for screening
 CC an anti-neoplastic agent, and can be used for producing a product which
 CC is the data collected with respect to the anti-neoplastic agent as a
 CC result of M1, and the data is sufficient to convey the chemical
 CC structure and/or properties of the agent. M1 can be used in the
 CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
 CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
 CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
 CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
 CC carcinoma, papillary carcinoma and Wilm's tumour.
 XX
 SQ Sequence 3370 BP; 689 A; 1002 C; 1005 G; 674 T; 0 other;
 Alignment Scores:
 Pred. No.: 0.698 Length: 3370
 Score: 12.00 Matches: 12
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.09% Indels: 0
 DB: Gaps: 0
 US-09-397-967-16 (1-1099) x ABL67952 (1-3370)
 OY 941 VALHISARGSPLEUALALALARGASMLEUVAL 952
 Db 2353 GTCCACGCGAGCTGCTGCCAGAAACATCTTGTCG 2388
 RESULT 75
 ID ABR49561 standard; DNA; 3415 BP.
 XX
 AC ABR49561;
 XX
 DT 15-JUL-2002 (first entry)
 XX
 DE Human cDNA 16658 encoding a novel kinase.
 XX
 KW Human; ss; gene; 16658; kinase; cellular proliferative disorder; cancer;
 KW carcinoma; tumour; adenocarcinoma; haematopoietic neoplastic disorder;
 KW leukaemia; lymphoma; brain disorder; cerebral ischaemia; infection;
 KW meningitis; brain abscess; acquire immunodeficiency syndrome; obesity;
 KW AIDS-related myopathy; prion disease; Alzheimer's disease; diabetes;
 KW Parkinson's disease; Huntington's disease; motor neurone disease;
 KW metabolic disorder; anorexia nervosa; pain; inflammation; ischaemia;
 KW irritable bowel syndrome; heart disorder; myocardial infarction;
 KW blood vessel disorder; atherosclerosis; bone metabolism disorder;
 KW osteoporosis; haematopoietic disorder; arthritis.
 XX
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FH 23..3415
 FT CDS
 FT /*tag= a
 FT /product= "kinase 16658"
 FT /note= "This coding sequence is specifically claimed
 FT in claim 1"
 XX
 PN WO200220800-A2.
 XX
 PD 14-MAR-2002.
 XX
 PF 03-AUG-2001; 2001WO-US24601.
 XX
 PR 01-SEP-2000; 2000US-229299P.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 XX

KW fatty liver; Gaucher's disease; Niemann-Pick's disease; anorectic;
 KW hypercholesterolaemia; obesity; gene therapy; cytostatic; anti-HIV;
 KW neuroprotective; hepatotropic; hypotensive; cardiac; nephrotropic;
 KW hyperlipidaemia; enzyme; gene; ss.
 OS Homo sapiens.
 XX
 XX
 FH Key Location/Qualifiers
 FT CDS 1..3345
 FT /*tag= a
 FT /product= "Human kinase (PKIN)-19"
 FT slq_peptide 1..84
 FT /*tag= b
 FT mat_peptide 85..3342
 FT /*tag= c
 FT /product= "Human mature kinase (PKIN)-19"
 FT
 FT
 PN WO200233099-A2;
 PD 25-APR-2002.
 PD 20-OCT-2001; 2001WO-US47728.
 PD 20-OCT-2000; 2000US-242410P.
 PR 27-OCT-2000; 2000US-244068P.
 PR 03-NOV-2000; 2000US-245708P.
 PR 09-NOV-2000; 2000US-247672P.
 PR 16-NOV-2000; 2000US-249565P.
 PR 22-NOV-2000; 2000US-252730P.
 PR 01-DEC-2000; 2000US-250807P.
 PR
 XX
 PA (INCYTE GENOMICS INC.
 XX
 PI Gururajan R, Baughn MR, Walla NK, Elliott VS, Xu Y, Arvizu C;
 PI Yao MG, Ramkumar J, Ding L, Tang YT, Hafalia AJA, Nguyen DB;
 PI Gandhi AR, Lu Y, Yue H, Burford N, Bandman O, Tribouley CM,
 PI Lal PG, Recipon SA, Lu DM, Borowsky ML, Thornton M, Swarnaker A;
 PI Thangavelu K, Khan FA, Ison CH;
 XX
 DR WPI: 2002-454603/48.
 DR P-PSDB: MAE24148.
 XX
 PT New human kinase polypeptide, for diagnosing, preventing and treating
 PT cancer, immune system disorders, growth and development disorders,
 PT cardiovascular disorders and lipid disorders
 PT
 XX
 PS Claim 5; Page 205..206; 210pp; English.
 XX
 CC The invention relates human kinases (PKIN) and their corresponding
 CC nucleic acid sequences. PKIN and its DNA are useful for diagnosing,
 CC treating and preventing cancer, an immune system disorder (e.g.,
 CC acquired immune deficiency syndrome (AIDS), Addison's disease, allergy,
 CC asthma, atherosclerosis, multiple sclerosis, psoriasis), disorders
 CC affecting growth and development (e.g., arteriosclerosis, cirrhosis,
 CC hepatitis), cardiovascular disorder (e.g., hypertension, myocardial
 CC infarction, Goodpasture's syndrome), and a lipid disorder (e.g., fatty
 CC liver, Gaucher's disease, Niemann-Pick's disease, hypercholesterolaemia,
 CC hyperlipidaemia, obesity), and for assessing the effects of exogenous
 CC compounds. Anti-PKIN antibody is useful in a diagnostic test for a
 CC condition or a disease associated with the expression of PKIN in a
 CC biological sample. A composition comprising PKIN or an agonist, or
 CC antagonist of PKIN is useful for treating a disease or condition
 CC associated with decreased or increased expression of functional PKIN.
 CC PKIN is useful in a number of drug screening techniques and to analyse
 CC the proteome of a tissue or cell type. PKIN DNA is useful for creating
 CC knockin humanised animals or transgenic animals to model human diseases,
 CC and in somatic or germline gene therapy. The present sequence is human
 CC PKIN cDNA.
 CC
 SO Sequence 3472 BP; 708 A; 1046 C; 1029 G; 689 T; 0 other;

Alignment Scores: 0.718 Length: 3472
 Pred. No.:
 ..

Score: 12.00 Matches: 12
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.09% Indels: 0
 DB: 24 Gaps: 0
 US-09-397-967-16 (1-1099) x AAQ38862 (1-3472)
 Oy 941 VALHISARGASPLEUAlAlaArgsnlleuVal 952
 Db 2611 GTTCATCGGACTTGGCCAGCAACAATCCTGGTA 2646
 RESULT 78
 ID AAQ90656 standard; cDNA; 3546 BP.
 AC AAQ90656;
 DT 11-NOV-1995 (first entry)
 XX Eph-related PTK Cek10 cDNA.
 DE
 XX
 KW Cek10; Eph; protein tyrosine-kinase; PTK; cancer; diagnosis;
 KW prognosis; ss.
 XX
 OS Gallus sp.
 XX
 FH Key Location/Qualifiers
 FT CDS 2..2923
 FT /*tag= a
 FT
 FT
 PN WO9515375-A.
 XX
 PD 08-JUN-1995.
 PD 07-SEP-1994; 94WO-US10140.
 PR 03-DEC-1993; 93US-0162809;
 PR
 XX
 PA (LJOL-) LA JOLLA CANCER RES FOUND.
 XX
 PI Pasquale EB, Sajjadi FG;
 XX
 DR WPI: 1995-215256/28.
 DR P-PSDB: AAR75708.
 XX
 PT Eph-related protein tyrosine kinase(s) - for monitoring and diagnosing
 PT cancer.
 PT
 XX
 PS Claim 2; Page 63-67; 129pp; English.
 XX
 CC Novel Eph-related PTK cDNA clone Cek10 (given in AAQ90656) and a
 CC variant clone, Cek10+ (AAQ90658), whose product contains a 15-amino
 CC acid insertion in the juxtamembrane domain, were isolated from a
 CC chick embryo library in lambda gcl1. Cek10 and cek10+ may
 CC originate from the same gene by alternative splicing. Cek10
 CC expression was prominent in the kidney, and to a lesser extent in
 CC the lung.
 CC
 SO Sequence 3546 BP; 742 A; 1100 C; 1031 G; 673 T; 0 other;

Alignment Scores: 0.732 Length: 3546
 Pred. No.: 12.00 Matches: 12
 Score: 12.00 Conservative: 0
 Percent Similarity: 100.00% Mismatches: 0
 Best Local Similarity: 100.00% Indels: 0
 Query Match: 1.09% Gaps: 0
 DB: 16
 US-09-397-967-16 (1-1099) x AAQ90656 (1-3546)
 Oy 941 VALHISARGASPLEUAlAlaArgsnlleuVal 952
 |||||||||||||||||||||||||||||||||||||||

DB 2189 GTGACCGAGACCTGCTGCCGCAACATCTGTC 2224

RESULT 79

ID ABK70012

ABK70012 standard; DNA: 3583 BP.

AC ABK70012;

DT 15-JUL-2002 (first entry)

DE cDNA encoding human Pro peptide #52.

Human; ss: gene; PRO; secreted protein; transmembrane protein;
genetic disorder; tumour; cancer.

Homo sapiens.

MO200224888-A2.

28-MAR-2002.

29-AUG-2001; 2001MO-US27099.

01-SEP-2000; 2000US-229896P.

05-SEP-2000; 2000US-230621P.

22-SEP-2000; 2000US-235147P.

10-NOV-2000; 2000MO-US30873.

12-JAN-2001; 2001US-261878P.

16-JAN-2001; 2001US-261910P.

16-JAN-2001; 2001US-262150P.

25-JAN-2001; 2001US-264395P.

02-FEB-2001; 2001US-266421P.

09-FEB-2001; 2001US-267623P.

28-FEB-2001; 2001MO-US06520.

09-MAR-2001; 2001US-274399P.

03-APR-2001; 2001US-280982P.

04-APR-2001; 2001US-282132P.

09-MAY-2001; 2001US-282189P.

25-MAY-2001; 2001MO-US17092.

01-JUN-2001; 2001MO-US17600.

20-JUN-2001; 2001MO-US19692.

29-JUN-2001; 2001MO-US21066.

09-JUL-2001; 2001MO-US21735.

(GETH) GENENTECH INC.

Baker KP, Eaton DL, Filvaroff E, Goddard A, Grimaldi JC;
Gurney AL, Smith V, Stephan J, Watanabe CK, Wood WI, Zhang Z;
Fong S;

WPI: 2002-362426/39.

P-PSDB: ABG34081.

New PRO polypeptides and polynucleotides encoding the polypeptides,
useful in gene therapy, chromosome identification, tissue typing, or
for genetic analysis of individuals with genetic disorders

Claim 2: Figure 103; 218pp; English.

This invention relates to the cDNA and protein sequences of novel
secreted and transmembrane polypeptides PRO polypeptides. The
invention also comprises a method for producing the proteins of the
invention by recombinant means and antibodies specific for the protein
proteins of the invention. The antibody may be used for detecting the PRO
polynucleotides may be used as hybridisation probes for a cDNA library
to isolate the full-length PRO cDNA or to isolate other cDNAs, to
construct hybridisation probes for mapping the gene which encodes that
PRO and for genetic analysis of individuals with genetic disorders. In
assays to identify other proteins or molecules involved in binding
reaction, to generate transgenic animals or knock-out animals which in

CC turn are useful in the development and screening of therapeutically
CC useful reagents, for chromosome identification, and tissue typing. The
CC PRO polypeptides are useful in gene therapy, and as molecular weight
CC markers for protein electrophoresis purposes. The sequences may
CC also be used to detect overexpression on PRO polypeptides in cancerous
CC tumours and for screening for differentially expressed genes using
CC microarray technology. The present sequence represents a cDNA encoding
CC a human PRO protein of the invention.

SQ Sequence 3583 BP; 1019 A; 816 C; 853 G; 895 T; 0 other:

Alignment Scores:

Pred. No.:	0.74	Length:	3583
Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.09%	Indels:	0
DB:	24	Gaps:	0

US-09-397-967-16 (1-1099) x ABK70012 (1-3583)

OY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952

DB 2633 GTTCATCGAGACCTACGCGTGGATATATCTGCTC 2668

RESULT 80

AAQ90658

ID AAQ90658 standard; cDNA: 3591 BP.

AC AAQ90658;

DT 11-NOV-1995 (first entry)

DE Eph-related PTK Cek10+ cDNA.

KW Cek10+; Eph; protein tyrosine-kinase; PTK; cancer; diagnosis;

KM prognosis; ss.

OS Gallus sp.

FT Key Location/Qualifiers

FT CDS 2..2968

FT /*tag= a

PN MO9515375-A.

PD 08-JUN-1995.

PF 07-SEP-1994; 94MO-US10140.

PR 03-DEC-1993; 93US-0162809.

PA (LJOL-) LA JOLLA CANCER RES FOUND.

PI Pasquale EB, Sajjadi FG;

PT WPI: 1995-215256/28.

DR P-PSDB: AAR75710.

XX Eph-related protein tyrosine kinase(s) - for monitoring and diagnosing
cancer.

Claim 2: Page 78-82; 129pp; English.

Novel Eph-related PTK cDNA clone Cek10 (given in AAQ90656) and a
variant clone, Cek10+ (AAQ90658), whose product contains a 15-amino
acid insertion in the juxtamembrane domain, were isolated from a
chick embryo library in lambda gt11. Cek10 and Cek10+ may
originate from the same gene by alternative splicing. Cek10
expression was prominent in the kidney, and to a lesser extent in
the lung.

Sequence 3591 BP; 753 A; 1112 C; 1042 G; 684 T; 0 other;

Alignment Scores:

Pred. No.:	0.741	Length:	3591
Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.09%	Indels:	0
DB:	16	Gaps:	0

US-09-397-967-16 (1-1099) x AAG90658 (1-3591)

OY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
 |||
 Db 2234 GTTCACGAGACCTGCGCCGCAACATCCTGTC 2269

RESULT 81
 AAV70208
 ID AAV70208 standard; DNA: 3592 BP.
 XX AAV70208;
 AC
 XX 11-FEB-1999 (first entry)
 DT
 XX
 DE Rat receptor tyrosine kinase Etk-2 encoding DNA.
 XX
 XX Receptor tyrosine kinase: Ror-1; Ror-2; Etk-1; Etk-2; detection;
 KW neurotrophin activity; TrkB; proto-oncogene; tyrosine kinase receptor;
 KW binding protein; BDNF; NT-3; diagnosis; ss.
 XX
 OS Rattus sp.
 XX
 FH Key Location/Qualifiers
 FT CDS 598..3444
 FT /*tag= a
 FT
 XX
 XX US5843749-A.
 PN
 XX
 PD 01-DEC-1998.
 XX
 XX 06-JUN-1995; 95US-0469537.
 PF
 XX
 PR 17-MAR-1995; 95US-0406247.
 PR 26-JUL-1991; 91US-0736559.
 PR 28-OCT-1993; 93US-0144992.
 PR 06-JUN-1995; 95US-0469537.
 XX
 XX (REG-) REGENERON PHARM INC.
 PA
 XX
 PI Maisongierre PC, Mastakowski P, Yancopoulos GD;
 XX
 DR WPI: 1999-044584/04.
 DR P-PSDB; AAM83148.
 XX
 XX DNA encoding receptor tyrosine kinase proteins - and corresponding
 PT proteins
 XX
 PS Claim 7; Fig 21; 194pp; English.
 XX
 CC The present invention describes nucleic acid molecules for Ror-1,
 CC Ror-2, Etk-1 and Etk-2. Also described are the corresponding proteins;
 CC Ror-1; Ror-2; Etk-1, and Etk-2. The proteins are orphan receptor
 CC tyrosine kinases. The present sequence encodes rat Etk-2.
 CC
 XX
 SO Sequence 3592 BP; 946 A; 916 C; 894 G; 834 T; 2 other;

Alignment Scores:

Pred. No.:	0.741	Length:	3592
Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.09%	Indels:	0
DB:	20	Gaps:	0

US-09-397-967-16 (1-1099) x AAV70208 (1-3592)

OY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
 |||
 Db 2977 GTTCATCGAGACCTAGCGCTAGGACATCTTGTC 3012

RESULT 82
 AAT72320
 ID AAT72320 standard; DNA: 3663 BP.
 XX
 XX AAT72320;
 AC
 XX 15-SEP-1997 (first entry)
 DT
 XX
 DE Embryonic stem cell kinase (Esk) DNA.
 XX
 XX Embryonic stem cell kinase; receptor tyrosine kinase; mucositis;
 KW epithelium; signal transduction; gene therapy; diagnosis; ss.
 XX
 OS Mus musculus.
 XX
 FH Key Location/Qualifiers
 FT CDS 355..3288
 FT /*tag= a
 FT sig_peptide 355..432
 FT /*tag= b
 FT mat_peptide 433..3285
 FT /*tag= c
 XX
 PN WO9723629-A1.
 XX
 PD 03-JUL-1997.
 XX
 XX 20-DEC-1996; 96WO-AU00826.
 PF
 XX
 PR 22-DEC-1995; 95AU-0007277.
 XX
 XX (AMRA-) AMRAD OPERATIONS PTY LTD.
 PA
 XX
 PI Boyd AM, Lickliter J;
 XX
 DR WPI: 1997-351065/32.
 DR P-PSDB; AAM19258.
 XX
 XX DNA encoding embryonic stem cell kinase, receptor tyrosine kinase -
 PT for production of modulators and antibodies, useful to treat
 PT mucositis and other disorders involving epithelium
 XX
 XX Claim 4; Page 48-54; 98pp; English.
 PS
 XX
 CC An isolated nucleic acid molecule (AAT72320) codes for novel mouse
 CC embryonic stem cell kinase (Esk), a receptor tyrosine kinase that
 CC belongs to the Eph subfamily. To isolate the molecule, RNA from
 CC the murine embryonic stem cell line W9.5 was subjected to RT-PCR
 CC using primers (see also AAT72321-24) based on conserved Eph
 CC sequences. A labelled probe from an isolated clone was used to
 CC screen a mouse liver cDNA library, identifying a single clone that
 CC included the complete coding region of Esk. The gene for murine
 CC Esk is located on bands B1 and B2 of chromosome 6. Esk nucleic
 CC acids can be used to produce recombinant Esk polypeptides, develop
 CC probes, and as sense and antisense molecules for diagnosis or
 CC therapy.
 CC
 XX
 SO Sequence 3663 BP; 869 A; 976 C; 1015 G; 801 T; 2 other;

Alignment Scores:

Pred. No.:	0.755	Length:	3663
Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.09%	Indels:	0
DB:	18	Gaps:	0

```

US-09-397-967-16 (1-1099) x AAT72320 (1-3663)
QY 941 VALHISARGAspleuAlaAlaArgAsnIleuVal 952
   |||||||
Db 2593 GTCCATAGAGACCTGGCTGCCAGAACATCTGTGTC 2628

RESULT 83
AA062461
ID AA062461 standard; cDNA; 3751 BP.
XX
AC AA062461;
XX
DT 09-NOV-1994 (first entry)
XX
DE Human embryonal kinase 2 receptor.
XX
KM Embryonal kinase; HEK; protein tyrosine kinase; PTK; tumour;
KM cancer; therapy; amplification; primer; polymerase chain reaction;
KM PCR; ss.
XX
OS Homo sapiens.
XX
FH Key 1.2973
FT CDS /tag= a
XX
PD DE4233782-A.
XX
PN 14-APR-1994.
XX
PF 07-OCT-1992; 92DE-4233782.
XX
PR 07-OCT-1992; 92DE-4233782.
XX
PS (CHEM-) CHEMOTHERAPEUTISCHES FORSCHUNG.
XX
PI Holtrich U, Ruebsamen-Waigmann H, Strebhardt K;
XX
DR WPI: 1994-127194/16.
XX
P-PSDB: AAR51899.
XX
PT Human embryonal kinase 2-receptor protein - useful in tumour
PT diagnosis and therapy
XX
PS Claim 4; Page 7-10; 11pp; German.
XX
CC RNA from human embryonic tissue was isolated. With the use of
CC primer P6(4) PTK-specific cDNA was synthesised. The cDNA was
CC amplified using primers P6(4) and N5. A 2097 bp DNA fragment was
CC obtained. Primers E3, P12 and E6 were then used in the isolation of
CC the C-terminal of the HEK2 receptor gene.
XX
SQ Sequence 3751 BP; 739 A; 1150 C; 1123 G; 739 T; 0 other;

Alignment Scores:
Pred. No.: 0.772 Length: 3751
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 15 Gaps: 0

US-09-397-967-16 (1-1099) x AA062461 (1-3751)
QY 941 VALHISARGAspleuAlaAlaArgAsnIleuVal 952
   |||||||
Db 2239 GTGCACCGGACCTGGCTGCTGCACACATCTGTGTC 2274

RESULT 84
AAA88548
ID AAA88548 standard; cDNA; 3768 BP.
XX
AC AAA88548;

```

```

XX
DT 22-JAN-2001 (first entry)
XX
DE Human CASB616 cDNA.
XX
KM CASB616; EPHB2; ERK; EPH3; EPH3T3; DRT; HEK5; EPHB2V;
KM receptor protein tyrosine kinase; human; antigen; colon cancer;
KM Ovary cancer; tumour; autoimmune disease; vaccine; gene therapy;
KM diagnosis; ss.
XX
OS Homo sapiens.
XX
FH Key 105..3068
FT CDS /tag= a
FT CDS /transl_except= (pos:2973..2975,aa:Leu)
XX
PN WO200053216-A2.
XX
PD 14-SEP-2000.
XX
PF 28-FEB-2000; 2000WO-EP01587.
XX
PR 05-MAR-1999; 99GB-0005124.
XX
PS (SMK) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI Vinals De Bassols XC;
XX
DR WPI: 2000-587384/55.
XX
P-PSDB: AAB19590.
XX
PT Vaccine composition for treating ovarian and colon cancer, comprises
PT CASB616 polypeptides, polynucleotides or antigen presenting cells
PT expressing the polypeptides
XX
PS Claim 3; Page 40-41; 57pp; English.
XX
CC The present sequence is that of cDNA coding for human CASB616 (see
CC AAB19590), a member of the EPH and EPH-related family of receptor
CC protein tyrosine kinases. CASB616 is also known as EPHB2, ERK,
CC EPH3, EPH3T3, DRT, HEK5 and EPHB2V. CASB616 polypeptides and
CC polynucleotides are important immunogens for specific prophylactic
CC or therapeutic immunization against tumours, especially colon
CC cancer (claimed) and ovarian cancer. They are specifically
CC expressed or highly over-expressed in tumours compared to normal
CC cells and can thus be targeted by antigen-specific immune
CC mechanisms leading to destruction of the tumour cells. They can
CC also be used to diagnose the occurrence of tumour cells. Their
CC inappropriate expression can also cause an induction of autoimmune
CC responses, which can be corrected through vaccination using the
CC CASB616 polypeptides or polynucleotides.
XX
SQ Sequence 3768 BP; 872 A; 1129 C; 1058 G; 709 T; 0 other;

Alignment Scores:
Pred. No.: 0.775 Length: 3768
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 21 Gaps: 0

US-09-397-967-16 (1-1099) x AAA88548 (1-3768)
QY 941 VALHISARGAspleuAlaAlaArgAsnIleuVal 952
   |||||||
Db 2334 GTTCACCGTGTGCTGCTGCCGCAACATCTGTGTC 2369

RESULT 85
AA090655
ID AA090655 standard; cDNA; 3776 BP.
XX
AC AA090655;

```

AC AA090655;
 XX
 DT 11-NOV-1995 (first entry)
 XX
 DE Eph-related PTK Cdk9 CDNA.
 XX
 KW Cdk9; Eph; protein tyrosine-kinase; PTK; cancer; diagnosis;
 KW prognosis; ss.
 XX
 OS Gallus sp.
 XX
 XX
 FH Key Location/Qualifiers
 FT CDS 290..311
 FT /*tag= a
 XX
 PN WO9515375-A.
 XX
 PD 08-JUN-1995.
 XX
 PF 07-SEP-1994; 94WO-US10140.
 XX
 PR 03-DEC-1993; 93US-0162809.
 XX
 PA (LJOL-) LA JOLLA CANCER RES FOUND.
 XX
 PI Pasquale EB, Sajjadi FG;
 XX
 DR WPI; 1995-215256/28.
 DR P-PSDB; AAR75707.
 XX
 PT Eph-related protein tyrosine kinase(s) - for monitoring and diagnosing
 PT cancer.
 XX
 PS Claim 2; Page 50-53; 129pp; English.
 XX
 CC Novel Eph-related PTK cDNA clone Cdk9 (given in AA090655) was
 CC isolated from a chick embryo library in lambda. Cdk9 protein
 CC (AAR75707) is closely related to Cdk5 (AAR75712). In adult tissues,
 CC Cdk9 expression is predominant in the thymus and detectable in
 CC brain, retina, kidney, lung and heart.
 CC
 XX
 SO Sequence 3776 BP; 1005 A; 881 C; 941 G; 949 T; 0 other;

Alignment Scores:
 Pred. No.: 0.777 Length: 3776
 Score: 12.00 Matches: 12
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.09% Indels: 0
 DB: 16 Gaps: 0

US-09-397-967-16 (1-1099) x AA090655 (1-3776)
 OY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
 DB 2477 GTGCATCGTATCTCGCAGCAGTAAACATCTTACTC 2512

RESULT 86
 ID ABA09147 standard: cDNA; 3836 BP.
 XX
 AC ABA09147;
 XX
 DT 11-JAN-2002 (first entry)
 XX
 DE Human protein Tyr kinase receptor homologue cDNA, SEQ ID NO:923.
 XX
 KW Human; cytokine; cell proliferation; cell differentiation; growth factor;
 KW haematopoiesis regulation; tissue growth; immunomodulator; activin;
 KW inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
 KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
 KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
 KW chronic inflammatory condition; proliferative retinopathy;

KW atherosclerosis; coronary heart disease; arterial ischaemia;
 KW bone disorder; osteoporosis; vascular growth disorder;
 KW tissue regeneration; wound healing; infection; immune disorder;
 KW cell culture; drug screening; gene therapy; antiinflammatory;
 KW antiasthmatic; antiarthritis; haemostatic; antiarteriosclerotic;
 KW cytostatic; osteopathic; vasotropic; cardiatic; antiviral; antibacterial;
 KW antifungal; vulnery; antitumor; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200157188-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 05-FEB-2001; 2001WO-US03800.
 XX
 PR 03-FEB-2000; 2000US-0496914.
 PR 27-APR-2000; 2000US-0560875.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Drmanac RT;
 XX
 DR WPI; 2001-457740/49.
 DR P-PSDB; ABB11903.
 XX
 PT Human proteins and DNA encoding sequences useful for preventing,
 PT treating or ameliorating a medical condition in a mammalian subject
 PT e.g. arthritis and cancer -
 XX
 XX
 PS Claim 1; Page 798-799; 1963pp; English.
 XX
 CC Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
 CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
 CC invention also relates to vectors and recombinant host cells comprising a
 CC nucleotide of the invention, methods of producing the novel polypeptides,
 CC antibodies against the polypeptides, methods of detecting the nucleotides
 CC or polypeptides in a sample, and methods of identifying compounds which
 CC bind to polypeptides of the invention. Although novel, many of the
 CC polypeptides of the invention have homology to known proteins, thereby
 CC giving an insight into their probable biological activities, and hence
 CC potential therapeutic applications. The polypeptides of the invention may
 CC have various activities, including cytokine, cell proliferation or cell
 CC differentiation activities; stem cell growth factor activity;
 CC haematopoiesis regulatory activity; tissue growth activity;
 CC immunomodulatory activity; activin- or inhibin-related activities;
 CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
 CC thrombolytic activities; receptor or ligand activities; or may be
 CC involved in oncogenesis, cancer cell proliferation or metastasis.
 CC Depending on their biological activities, polypeptides and nucleotides of
 CC the invention are useful for preventing, treating or ameliorating medical
 CC conditions, e.g., by protein or gene therapy. Such conditions include
 CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
 CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
 CC proliferative retinopathy, atherosclerosis, coronary heart disease,
 CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
 CC vascular growth. Polypeptides involved with tissue regeneration and
 CC repair (or nucleic acids encoding them) may be used to promote wound
 CC healing (e.g., of burns, incisions and ulcers), while those with
 CC immunomodulatory activities may be used in the treatment of viral,
 CC bacterial and fungal infections in addition to immune disorders.
 CC Polypeptides with growth factor activity may be used in cell cultures to
 CC promote cell growth. For example, such polypeptides may be used to
 CC manipulate stem cells in culture to give rise to neuroepithelial cells
 CC that can be used to augment or replace cells damaged by illness,
 CC autoimmune disease or accidental damage. The polypeptides and nucleotides
 CC may also be used in the diagnosis of the above conditions, and in drug
 CC screening techniques. The present sequence represents a cDNA encoding a
 CC novel human polypeptide of the invention.
 XX
 SO Sequence 3836 BP; 759 A; 1190 C; 1140 G; 746 T; 1 other;

Alignment Scores:

Pred. No.: 0.788 Length: 3836
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservatve: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 22 Gaps: 0

US-09-397-967-16 (1-1099) x ABA09147 (1-3836)

OY 941 VALHISARGASPLEUALAALARGASNTLEUVAL 952
Db 2290 GTGCACCGACCTGCTGCGACAACATCTTGTG 2325

RESULT 87
AAD27896
ID AAD27896 standard; CDNA: 3945 BP.
XX
AC AAD27896;
XX
DT 31-MAY-2002 (first entry)
XX
DE Human EphB4 CDNA.
XX
KW Artery; vein; endothelial cell; obstructed blood vessel; coronary artery;
KW atherosclerosis; graft; gene therapy; human; EphB4; arterial marker; ss.
XX
OS Homo sapiens.
XX
PN WO200211785-A2.
XX
PD 14-FEB-2002.
XX
PF 03-AUG-2001; 2001WO-US24405.
XX
PR 03-AUG-2000; 2000US-222759P.
XX
PA (UTAH) UNIV UTAH.
XX
PI L1 DY;
XX
DR WPI: 2002-217155/27.
XX
PT Inducing arterial morphology in vein by contacting endothelial cells in
PT vein to a polynucleotide encoding gene capable of inducing endothelial
PT remodeling to transfer the polynucleotide into the cells

PS Claim 6; Page 30-34; 39pp; English.

CC The invention relates to a method of inducing arterial morphology in a
CC vein, comprising contacting endothelial cells in the vein to at least
CC one polynucleotide encoding a gene that is capable of inducing
CC endothelial remodeling for a time sufficient to transfer the
CC polynucleotide into the endothelial cells. A method for treating a
CC patient having an obstructed blood vessel, e.g. coronary artery
CC affected by atherosclerosis, is also provided which comprises providing
CC a graft comprising endothelial cells, contacting the endothelial cells of
CC the graft to the polynucleotide, removing a section of the obstructed
CC blood vessel, and grafting the graft in place of the removed section of
CC the obstructed blood vessel, where a graft is provided by harvesting a
CC section of a vein from the patient. The present sequence is human
CC EphB4 CDNA used in the invention for endothelial remodeling.
CC The EphB4 gene encodes an arterial specific molecular marker.
XX
SQ Sequence 3945 BP; 766 A; 1186 C; 1212 G; 781 T; 0 other;

Alignment Scores:
Pred. No.: 0.809 Length: 3945
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservatve: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 24 Gaps: 0

US-09-397-967-16 (1-1099) x AAD27896 (1-3945)

OY 941 VALHISARGASPLEUALAALARGASNTLEUVAL 952
Db 2294 GTGCACCGACCTGCTGCGACAACATCTTGTG 2329

RESULT 88
ABL68537
ID ABL68537 standard; DNA: 3945 BP.
XX
AC ABL68537;
XX
DT 15-MAY-2002 (first entry)
XX
DE Kidney cancer related gene sequence SEQ ID NO:6874.
XX
DE Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilms' tumour; adenocarcinoma;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO200194629-A2.
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US10838.
XX
PR 05-JUN-2000; 2000US-209473P.
XX
PR 05-JUN-2000; 2000US-209531P.
XX
PR 18-SEP-2000; 2000US-233133P.
XX
PR 18-SEP-2000; 2000US-233617P.
XX
PR 20-SEP-2000; 2000US-234009P.
XX
PR 20-SEP-2000; 2000US-234034P.
XX
PR 20-SEP-2000; 2000US-234052P.
XX
PR 22-SEP-2000; 2000US-234509P.
XX
PR 22-SEP-2000; 2000US-234567P.
XX
PR 25-SEP-2000; 2000US-234923P.
XX
PR 25-SEP-2000; 2000US-234924P.
XX
PR 25-SEP-2000; 2000US-235077P.
XX
PR 25-SEP-2000; 2000US-235082P.
XX
PR 25-SEP-2000; 2000US-235134P.
XX
PR 25-SEP-2000; 2000US-235280P.
XX
PR 26-SEP-2000; 2000US-235637P.
XX
PR 26-SEP-2000; 2000US-235638P.
XX
PR 27-SEP-2000; 2000US-235711P.
XX
PR 27-SEP-2000; 2000US-235720P.
XX
PR 27-SEP-2000; 2000US-235840P.
XX
PR 27-SEP-2000; 2000US-235863P.
XX
PR 28-SEP-2000; 2000US-236028P.
XX
PR 28-SEP-2000; 2000US-236032P.
XX
PR 28-SEP-2000; 2000US-236033P.
XX
PR 28-SEP-2000; 2000US-236034P.
XX
PR 28-SEP-2000; 2000US-236109P.
XX
PR 28-SEP-2000; 2000US-236111P.
XX
PR 29-SEP-2000; 2000US-236842P.
XX
PR 29-SEP-2000; 2000US-236891P.
XX
PR 02-OCT-2000; 2000US-237172P.
XX
PR 02-OCT-2000; 2000US-237173P.
XX
PR 02-OCT-2000; 2000US-237278P.
XX
PR 02-OCT-2000; 2000US-237294P.
XX
PR 02-OCT-2000; 2000US-237295P.
XX
PR 02-OCT-2000; 2000US-237316P.
XX
PR 03-OCT-2000; 2000US-237425P.
XX
PR 03-OCT-2000; 2000US-237598P.
XX
PR 03-OCT-2000; 2000US-237604P.
XX
PR 03-OCT-2000; 2000US-237606P.
XX
PR 03-OCT-2000; 2000US-237608P.
XX
PR 01-NOV-2000; 2000US-244667P.
XX
PR 01-NOV-2000; 2000US-245084P.
XX
PA (AVAL-) AVALON PHARM.

```

XX  Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI  Soppet DR, Weaver Z;
XX
XX  WPI: 2002-188264/24.
XX
PT  Screening for anti-neoplastic agent involves exposing cells to a
PT  chemical agent to be tested for anti-neoplastic activity, and
PT  determining a change in expression of a gene of a signature gene set
XX
PS  Claim 1; SEQ ID 6874; 44pp; English.
XX
XX  The present invention describes a method (M1) for screening for an
XX  anti-neoplastic agent. The method involves exposing cells to a chemical
XX  agent to be tested for anti-neoplastic activity, determining a change in
XX  expression of at least one gene (I) of a signature gene set, where (i)
XX  comprises a sequence (S) selected from 8447 sequences (given in ABU61664
XX  to ABU70110), or is at least 95% identical to (S), where a change in
XX  expression is indicative of anti-neoplastic activity. (i) has cytostatic
XX  activity and can be used in gene therapy. M1 can be used for screening
XX  an anti-neoplastic agent, and can be used for producing a product which
XX  is the data collected with respect to the anti-neoplastic agent as a
XX  result of M1, and the data is sufficient to convey the chemical
XX  structure and/or properties of the agent. M1 can be used in the
XX  treatment of cancer such as colon, breast, stomach, lung, thyroid,
XX  oesophagagal, ovarian, kidney, prostate or pancreatic cancer,
XX  adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
XX  infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
XX  carcinoma, papillary carcinoma and Wilms' tumor.
XX
SQ  Sequence 3945 BP; 763 A; 1188 C; 1215 G; 779 T; 0 other;

Alignment Scores:
Pred. No.:      0.809      Length:      3945
Score:          12.00      Matches:      12
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches:  0
Query Match:    1.09%      Indels:      0
DB:             24         Gaps:        0

US-09-397-967-16 (1-1099) x ABU68537 (1-3945)

OY  941 VALHISARGASPLEUALAALARGASNLLEUVAL 952
    |||||||
DB  2294 GTCACCGAGACGTGGCTGCGCAACATCCTAGTC 2329

RESULT 89
ID  AAA88549 standard; cDNA: 3949 BP.
XX
XX  AAA88549;
XX
XX  22-JAN-2001 (first entry)
XX
DE  Human CASB616 cDNA.
XX
XX  CASB616; EPHB2; ERK; EPH3; EPH3; DRT; HEK5; EPHB2V;
XX  receptor protein tyrosine kinase; human; colon cancer;
XX  ovary cancer; tumour; autoimmune disease; vaccine; gene therapy;
XX  diagnosis; ss.
XX
OS  Homo sapiens.
XX
XX  Key Location/Qualifiers
XX  CDS 26..3193
XX  FT /*tag= a
XX
XX  WO200053216-A2.
XX
XX  14-SEP-2000.
XX
XX  28-FEB-2000; 2000WO-EP01587.
XX

```

```

PR  05-MAR-1999; 99GB-0005124.
XX
XX  (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
XX  Vinals De Bassols YC;
XX
XX  WPI: 2000-587384/55.
XX
XX  P-PSDB; AAB19591.
XX
PT  Vaccine composition for treating ovarian and colon cancer, comprises
PT  CASB616 polypeptides, polynucleotides or antigen presenting cells
PT  expressing the polypeptides
XX
XX
PS  Claim 3; Page 41-42; 57pp; English.
XX
XX  The present sequence is that of cDNA coding for human CASB616 (see
XX  AAB19591), a member of the EPH and EPH-related family of receptor
XX  protein tyrosine kinases. CASB616 is also known as EPHB2, ERK,
XX  EPH3, EPH3, DRT, HEK5 and EPHB2V. CASB616 polypeptides and
XX  polynucleotides are important immunogens for specific prophylactic
XX  or therapeutic immunization against tumours, especially colon
XX  cancer (claimed) and ovarian cancer. They are specifically
XX  expressed or highly over-expressed in tumours compared to normal
XX  cells and can thus be targeted by antigen-specific immune
XX  mechanisms leading to destruction of the tumour cells. They can
XX  also be used to diagnose the occurrence of tumour cells. Their
XX  inappropriate expression can also cause an induction of autoimmune
XX  responses, which can be corrected through vaccination using the
XX  CASB616 polypeptides or polynucleotides.
XX
SQ  Sequence 3949 BP; 934 A; 1142 C; 1105 G; 768 T; 0 other;

Alignment Scores:
Pred. No.:      0.81      Length:      3949
Score:          12.00      Matches:      12
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches:  0
Query Match:    1.09%      Indels:      0
DB:             21         Gaps:        0

US-09-397-967-16 (1-1099) x AAA88549 (1-3949)

OY  941 VALHISARGASPLEUALAALARGASNLLEUVAL 952
    |||||||
DB  2252 GTTCACCGAGACGTGGCTGCGCAACATCCTAGTC 2287

RESULT 90
ID  AAQ49757 standard; DNA: 3969 BP.
XX
XX  AAQ49757;
XX
XX  10-MAR-1994 (first entry)
XX
DE  PTK gene HPTK5.
XX
XX  PTK; protein tyrosine kinase; catalytic domain; c-kit; hepatoma cell;
XX  amplification; primer; polymerase chain reaction; PCR; ds.
XX
OS  Homo sapiens.
XX
XX  Key Location/Qualifiers
XX  CDS 89..3969
XX  FT /*tag= a
XX  FT /note= "the first stop codon is at position 3051-3053"
XX  FT m1sc-RNA 701..800
XX  FT /*tag= b
XX  FT /note= "100 bases are missing from the sequence
XX  of the specification and entered
XX  as (N)100 by the indexer"
XX
XX  WO9315201-A.
XX

```

PD 05-AUG-1993.
 XX
 PF 22-JAN-1993; 93WO-US00586.
 XX
 PR 22-JAN-1992; 92US-0826935.
 XX
 PA (NENE-) NEW ENGLAND DEACONESS HOSPITAL.
 XX
 PI Avraham H, Cowley S, Groopman J, Scadden D;
 XX
 DR WPI; 1993-320330/40.
 DR P-PSDB; AAR41896.
 XX
 PT New protein tyrosine kinase genes and proteins encoded by genes -
 PT are of human mega-karyocytic origin
 XX
 PS Claim 2; fig 8; 60pp; English.
 XX
 CC PTK genes were identified using two sets of degenerative
 CC oligonucleotide primers: a first set which amplifies all PTK DNA
 CC segments (AAQ49743-44), and a second set which amplifies highly
 CC conserved sequences present in the catalytic domain of the c-kit
 CC subgroup of PTKs (AAQ49745-46). The PTK genes identified are described
 CC in AAQ49747-57 and AAR41897-02.
 CC The Hprt5 gene is expressed in human hepatoma cells.
 XX
 SQ Sequence 3969 BP; 777 A; 1149 C; 1180 G; 763 T; 100 other;
 Alignment Scores:
 Pred. No.: 0.814 Length: 3969
 Score: 12.00 Matches: 12
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.09% Indels: 0
 DB: 14 Gaps: 0
 US-09-397-967-16 (1-1099) x AAQ49757 (1-3969)
 QY 941 VALHISARGASPLEUALAALARGASNIILEUVAL 952
 DB 2298 GTCCACCGAGACCTGCTGCTCGCACATCTTAGTC 2333
 RESULT 91
 AAT03099
 ID AAT03099 standard; DNA; 3969 BP.
 AC AAT03099;
 XX
 DT 14-FEB-1996 (first entry)
 XX
 DE Protein tyrosine-kinase Hprt5 gene.
 XX
 KW Protein tyrosine-kinase; PTK; Hprt5; hepatoma transmembrane kinase;
 KW agonist; cell growth; differentiation; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 90..3053
 FT sig_peptide /tag= a
 FT 90..134
 FT mat_peptide /tag= b
 FT 135..3050
 FT /tag= c
 FT /product= Hprt5
 XX
 PN WO9527061-A1.
 XX
 PD 12-OCT-1995.
 XX
 PF 04-APR-1995; 95WO-US04228.
 XX
 PR 04-APR-1994; 94US-0222616.

XX
 PA (GETH) GENENTECH INC.
 XX
 PI Bennett BD, Goeddel D, Lee JM, Matthews W, Tsai SP;
 PI Wood WI;
 XX
 DR WPI; 1995-36160/47.
 DR P-PSDB; AAR85930.
 XX
 PT Agonist antibodies which activate specific protein tyrosine
 PT kinase(s) - also activate chimeric proteins of kinase extracellular
 PT domain and Ig constant domain, useful for studying, and therapeutic
 PT modulation of, cell growth and differentiation
 XX
 PS Disclosure; Page 62-67; 125pp; English.
 XX
 CC DNA probes based on protein tyrosine-kinase (PTK) sequences were used
 CC to screen cDNA libraries to identify novel PTK genes. The Hprt5
 CC gene (AAT03099) was expressed in human hepatoma cells and encoded a
 CC protein (AAR85930) showing homology to known PTKs. The gene was used to
 CC construct an Hprt5 extracellular domain-IgG FC fusion protein.
 XX
 SQ Sequence 3969 BP; 786 A; 1187 C; 1214 G; 782 T; 0 other;
 Alignment Scores:
 Pred. No.: 0.814 Length: 3969
 Score: 12.00 Matches: 12
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.09% Indels: 0
 DB: 16 Gaps: 0
 US-09-397-967-16 (1-1099) x AAT03099 (1-3969)
 QY 941 VALHISARGASPLEUALAALARGASNIILEUVAL 952
 DB 2298 GTCCACCGAGACCTGCTGCTCGCACATCTTAGTC 2333
 RESULT 92
 AAO90972
 ID AAO90972 standard; cDNA to mRNA; 4027 BP,
 AC AAO90972;
 XX
 DT 24-NOV-1995 (first entry)
 XX
 DE Protein p140 cDNA from rat skeletal muscle myoblast cell line L6.
 XX
 KW Protein p140; insulin; tyrosine phosphorylation; ss.
 XX
 OS Rattus rattus.
 XX
 PN EP659883-A.
 XX
 PD 28-JUN-1995.
 XX
 FE 24-NOV-1994; 94EP-0118524.
 XX
 PR 24-NOV-1993; 93JP-0315806.
 XX
 PA (ONOV) ONO PHARM CO LTD.
 XX
 PI Kitagawa K, Ohno H, Tajima H;
 XX
 DR WPI; 1995-226291/30.
 XX
 PT Isolated protein p140 polypeptide - and treatment of diabetes based
 PT on tyrosine phosphorylation of protein p140.
 XX
 PS Claim 5; Page 25-27; 42pp; English.
 CC
 CC A cDNA library was established from rat skeletal myoblast cell line
 CC L6. DNA fragments of approx. 400 bp were recovered and subjected to

CC cloning. 20 plasmids were sequenced. Sequence data of cDNA fragments
 CC were constructed to the linkage sequences with the DNA programme
 CC DNASTIS. The basic sequence portrayed in AAQ90972 was hence constructed.
 CC From sequence data of the whole cDNA length, the ORF was determined.
 CC The AA sequence was further translated and the sequence thus
 CC established is illustrated in AAR75843. One of the frames possesses
 CC the 2993-bp ORF that was approximated to 3000 bp of the whole ORF
 CC length of the ECK family.

XX Sequence 4027 BP; 885 A; 1141 C; 1166 G; 835 T; 0 other;
 SQ

Alignment Scores:

Pred. No.:	0.825	Length:	4027
Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.09%	Indels:	0
DB:	16	Gaps:	0

US-09-397-967-16 (1-1099) x AAQ90972 (1-4027)

QY 941 VALHISARGASPLEUALAALARGASNIIEUVAL 952
 |||||||||||||||||||||||||||||||||||
 DB 2509 GTGCACCGTGACCTGCTGCCGCAACATCTTGTG 2544

RESULT 93

AAQ90982
 ID AAQ90982 standard; cDNA to mRNA; 4027 BP.

XX AAQ90982;

XX 24-NOV-1995 (first entry)

XX Protein p140 cDNA from rat skeletal muscle myoblast cell line L6.

XX Protein p140; insulin; tyrosine phosphorylation; ss.

XX Rattus rattus.

XX Key Location/Qualifiers
 FT CDS 262..3243
 FT /*tag= a

XX EP659883-A.

XX 28-JUN-1995.

XX 24-NOV-1994; 94EP-0118524.

XX 24-NOV-1993; 93JP-0315806.

XX (ONOV) ONO PHARM CO LTD.

XX Kitagawa K, Ohno H, Tajima H;

XX WPI: 1995-226291/30.

XX P-PSDB; AAR75844.

XX Isolated protein p140 polypeptide - and treatment of diabetes based
 PT on tyrosine phosphorylation of protein p140.

XX Example; Page 28-33; 42pp; English.

XX AAQ90982 does not appear to be referred to in the spec. although
 CC it is listed in the sequence listings (SEQ ID No. 4). It is the
 CC same length as SEQ ID No. 3 and may be identical to it (see
 CC AAQ90972).

XX Sequence 4027 BP; 885 A; 1141 C; 1166 G; 835 T; 0 other;
 SQ

Alignment Scores:

Pred. No.:	0.825	Length:	4027
Score:	12.00	Matches:	12

Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.09%	Indels:	0
DB:	16	Gaps:	0

US-09-397-967-16 (1-1099) x AAQ90982 (1-4027)

QY 941 VALHISARGASPLEUALAALARGASNIIEUVAL 952
 |||||||||||||||||||||||||||||||||||
 DB 2509 GTGCACCGTGACCTGCTGCCGCAACATCTTGTG 2544

RESULT 94

AAQ90660
 ID AAQ90660 standard; DNA; 4049 BP.

XX AAQ90660;

XX 11-NOV-1995 (first entry)

XX Eph-related PTK Cdk5.

XX Cdk5; Eph; protein tyrosine-kinase; PTK; cancer; diagnosis;

XX prognosis; ss.

XX Gallus sp.

XX Key Location/Qualifiers
 FT CDS 10..2997
 FT /*tag= a

XX WO9515375-A.

XX 08-JUN-1995.

XX 07-SEP-1994; 94WO-US101140.

XX 03-DEC-1993; 93US-0162809.

XX (LJOL-) LA JOLLA CANCER RES FOUND.

XX Pasquale EB, Sajjadi EG;

XX WPI: 1995-215256/28.

XX P-PSDB; AAR75712.

XX Eph-related protein tyrosine kinase(s) - for monitoring and diagnosing
 PT cancer.

XX Disclosure; Page 92-96; 129pp; English.

XX Probes derived from the Eph-related PTKs Cdk4 (AAQ90659) and Cdk5
 CC (AAQ90660) were used to isolate novel cDNA clones (AAQ90652-58,
 CC AAQ90661-62) from chicken embryo and embryonic brain libraries.

XX Sequence 4049 BP; 1010 A; 1011 C; 1072 G; 956 T; 0 other;
 SQ

Alignment Scores:

Pred. No.:	0.829	Length:	4049
Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.09%	Indels:	0
DB:	16	Gaps:	0

US-09-397-967-16 (1-1099) x AAQ90660 (1-4049)

QY 941 VALHISARGASPLEUALAALARGASNIIEUVAL 952
 |||||||||||||||||||||||||||||||||||
 DB 2263 GTGCACCGTGACCTGCTGCCGCAACATCTTGTG 2298

RESULT 95

AAQ90657
 ID AAQ90657 standard; DNA; 4097 BP.


```

XX AC AA090657;
XX
XX 11-NOV-1995 (first entry)
XX DT
XX DE Eph-related PTK Cdk5+.
XX KM Cdk5+; Eph; protein tyrosine-kinase; PTK; cancer; diagnosis;
XX KM prognosis; ss.
XX OS
XX Gallus sp.
XX FT
XX FT Key Location/Qualifiers
XX CDS 10..3045
XX /*tag= a
XX
XX PN W09515375-A.
XX
XX PD 08-JUN-1995.
XX
XX PF 07-SEP-1994; 94WO-US10140.
XX
XX PR 03-DEC-1993; 93US-0162809.
XX
XX PA (LJOL-) LA JOLLA CANCER RES FOUND.
XX
XX PI Pasquale EB, Sajjadi FG;
XX
XX DR WPI: 1995-215256/28.
XX DR P-PSDB: AAR75709.
XX
XX PT Eph-related protein tyrosine kinase(s) - for monitoring and diagnosing
XX PT cancer.
XX PS Claim 2; Page 71-75; 129pp; English.
XX
XX CC A cDNA clone encoding a novel variant of EPH-related PTK Cdk5,
XX CC Cdk5+ (AA090657), was isolated from a chick embryo library in
XX CC lambda gt11. Cdk5+ protein (AAR75709) contains a 16-amino acid
XX CC insertion in the juxtamembrane domain, and be a result of
XX CC alternative splicing. Cdk5+ is exclusively expressed in the CNS.
XX
XX SO Sequence 4097 BP; 1014 A; 1029 C; 1087 G; 967 T; 0 other;

Alignment Scores:
Pred. No.: 0.839 Length: 4097
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 16 Gaps: 0

US-09-397-967-16 (1-1099) x AA090657 (1-4097)
OY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
DB 2311 GTGCACCGGGACCTGCTGCCCGCAACATTCGTGTC 2346

RESULT 96
AA053471
ID AA053471 standard; CDNA: 4281 BP.
XX
XX AC AA053471;
XX
XX DT 16-JUN-1994 (first entry)
XX
XX DE elk CDNA.
XX
XX KW Lambda gt11; expression vector; lambda-BI-Elk; protein tyrosine kinase;
XX KW Elk; B1; Eph; subfamily; receptor-like tyrosine kinase; eph;
XX KW phosphorylation; phosphorylated kinase insert domain; growth factor;
XX KW receptor kinase; platelet-derived growth factor receptor; ss.
XX

```

```

OS Rattus rattus.
XX
XX FH Key Location/Qualifiers
XX CDS 367..3321
XX FT /*tag= a
XX FT /product= Elk
XX
XX PN CA2083521-A.
XX
XX PD 01-OCT-1993.
XX
XX PF 23-NOV-1992; 92CA-2083521.
XX
XX PR 31-MAR-1992; 92US-0861390.
XX
XX PA (MOUN ) MOUNT SINAI HOSPITAL CORP.
XX
XX PI Letwin K, Pawson A, Reedijk M;
XX
XX DR WPI: 1993-406300/51.
XX DR P-PSDB: AAR44513.
XX
XX PT Expression of phosphorylated exogenous protein - in host cells
XX PT transformed with two vectors, one for the protein, the other for
XX PT catalytic domain of protein kinase
XX
XX PS Disclosure: Fig 3; 55pp; English.
XX
XX CC This sequence represents the elk cDNA which encodes the protein
XX CC tyrosine kinase, Elk. The Elk gene, B1, encode a protein which is
XX CC a member of the Eph subfamily of protein tyrosine kinases. The Elk
XX CC product is very similar to two other receptor-like tyrosine kinases,
XX CC eph and eck. Lambda-BI-Elk may be used in the production of
XX CC phosphorylated exogenous protein along with a further vector encoding
XX CC the desired exogenous protein. These plasmid may be used to produce
XX CC phosphorylated proteins in host cells which have no intrinsic capacity
XX CC for phosphorylation, eg. bacteria. The system may be used for the
XX CC expression of the phosphorylated kinase insert domain of a growth
XX CC factor receptor kinase eg. platelet-derived growth factor receptor.
XX
XX SO Sequence 4281 BP; 1075 A; 1170 C; 1121 G; 915 T; 0 other;

Alignment Scores:
Pred. No.: 0.874 Length: 4281
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 14 Gaps: 0

US-09-397-967-16 (1-1099) x AA053471 (1-4281)
OY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
DB 2587 GTGCACCGGGACCTGCTGTAAGAACATTCGTGTC 2622

RESULT 97
AA092641
ID AA092641 standard; CDNA to mRNA: 4290 BP.
XX
XX AC AA092641;
XX
XX DT 06-MAR-1996 (first entry)
XX
XX DE Human non-differentiated blood cell tyrosine kinase receptor.
XX
XX KW Human; tyrosine kinase; receptor; non-differentiated; blood cell;
XX KW detection; cancer; bone marrow; transplant; treatment; ds.
XX
XX OS Homo sapiens.
XX
XX FH Key Location/Qualifiers
XX CDS 410..3373

```

```

FT      sig_peptide      /tag= a
FT      410..454
FT      /tag= b
FT      mat_peptide      455..3370
FT      /tag= c
XX      W09515386-A1.
XX      08-JUN-1995.
XX      02-DEC-1994; 94MO-JP02035.
XX      02-DEC-1993; 93JP-0302704.
XX      (ASAH ) ASAH KASEI KOGYO KK.
XX      Oono M, Sakano S;
XX      WPI: 1995-215266/28.
XX      P-PSDB: AAR89263, AAR76468.
XX      Polypeptide and modified versions having a receptor tyrosine kinase
XX      activity - are expressed in a non-differentiated blood cell but
XX      undergo a reduction in the expression level as the
XX      non-differentiated cell differentiates.
XX      Claim 13; Pages 55-62; 73pp; Japanese.
XX      AA092641 encodes AAR89263 human non-differentiated blood cell tyrosine
XX      kinase receptor (hTKR). As the blood cell differentiates the level
XX      of hTKR undergoes a reduction. An anti-hTKR antibody can be used to
XX      screen for non-differentiated blood cells in a biological sample,
XX      useful in the detection and treatment of diseased blood cells, e.g.
XX      cancerous cells etc., and esp. in bone marrow transplantation.
XX      SQ Sequence 4290 BP; 824 A; 1310 C; 1342 G; 809 T; 5 other;

Alignment Scores:
Pred. No.: 0.876 Length: 4290
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 16 Gaps: 0

US-09-397-967-16 (1-1099) x AA092641 (1-4290)
QY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
Db 2618 GTCACCGAGACCTGGCTGCTGCGACATCCTAGTC 2653

RESULT 98
AAT42593
ID AAT42593 standard; DNA; 4290 BP.
XX AAT42593;
XX AC AAT42593;
XX DT 17-JAN-1997 (first entry)
XX DE Coding sequence for receptor-type tyrosine kinase protein.
XX KW Receptor-type tyrosine kinase; ligand; coomassie staining; PAS staining;
XX human; ss.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX misc_difference 54..59
XX /tag= a
XX /note= "unspecified nucleotides, represented in
XX CDS 410..3373
XX /tag= b
XX /tag= b

```

```

FT      sig_peptide      /product= receptor-type tyrosine kinase protein
FT      410..454
FT      /tag= c
FT      mat_peptide      455..3370
FT      /tag= d
XX      JP08188596-A.
XX      23-JUL-1996.
XX      13-JAN-1995; 95JP-0003677.
XX      09-NOV-1994; 94JP-0275411.
XX      19-OCT-1994; 94JP-0253848.
XX      (ASAH ) ASAH KASEI KOGYO KK.
XX      WPI: 1996-388601/39.
XX      P-PSDB: AAM06335.
XX      New ligand for receptor type tyrosine kinase - has mol.wt. 22-25
XX      kilo(dalton(s) and is positive for Coomassie and PAS staining
XX      Example 16; Page 39-43; 51pp; Japanese.
XX      This sequence represents the coding sequence for the full length
XX      receptor-type tyrosine kinase protein. The encoded protein sequence is
XX      recognised by a protein of the invention. The proteins of the invention
XX      have a molecular weight of 23500 (plus or minus 1500) Da, and are
XX      positive for Coomassie staining and PAS staining. The N-terminus of two
XX      proteins of the invention (see AAM06333 and AAM06334) is represented by
XX      AAM06332. The proteins of the invention bind to the fragment of the
XX      encoded protein represented by AAM06330. The protein is a new ligand of
XX      receptor-type tyrosine kinases, and can be prepared by standard
XX      recombinant techniques.
XX      SQ Sequence 4290 BP; 824 A; 1310 C; 1342 G; 809 T; 5 other;

Alignment Scores:
Pred. No.: 0.876 Length: 4290
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 17 Gaps: 0

US-09-397-967-16 (1-1099) x AAT42593 (1-4290)
QY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
Db 2618 GTCACCGAGACCTGGCTGCTGCGACATCCTAGTC 2653

RESULT 99
AAT18394
ID AAT18394 standard; cDNA to mRNA; 4290 BP.
XX AAT18394;
XX AC AAT18394;
XX DT 18-OCT-1996 (first entry)
XX DE Receptor type tyrosine kinase (TK) coding sequence.
XX KW Receptor type tyrosine kinase; TK; ligand; differentiation;
XX haematopoietic stem cell; tyrosine; bone marrow; leukaemia; ss.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX CDS 410..3373
XX /tag= a
XX /product= Receptor type tyrosine kinase.
XX sig_peptide 410..454
XX /tag= b
XX /tag= b

```

```

FT mat-peptide 455..3370
FT /*tag= c
PN W09611212-A1.
PD 18-APR-1996.
XX
XX 09-OCT-1995; 95WO-JP02069.
XX
XX 22-DEC-1994; 94JP-0320712.
XX 07-OCT-1994; 94JP-0244433.
XX 26-OCT-1994; 94JP-0262882.
XX
XX (ASAH ) ASAH KASEI KOGYO KK.
XX
XX Ohno M, Sakano S;
XX
XX WPI: 1996-209809/21.
XX P-PSDB: AAR94652, AAR94653.
XX
XX Lysand peptide binding to receptor-type tyrosine kinase - enhances
XX intracellular tyrosine phosphorylation, useful for investigation of
XX undifferentiated blood cell behaviour
XX
XX Claim 1; Page 154-161; 193pp; Japanese.
XX
XX A lysand polypeptide which binds to the extracellular part of a
XX specific receptor-type tyrosine kinase and induces phosphorylation
XX of tyrosine within the cell can be used in the study of the
XX differentiation of blood cells such as the haematopoietic stem
XX cells; of disease processes such as leukemia, and of the biology of
XX bone marrow transplantation. The lysand plays a role in the
XX differentiation process and the specific lysand target is expressed
XX in undifferentiated blood cells.
XX
SQ Sequence 4290 BP; 824 A; 1323 C; 1329 G; 809 T; 5 other;

Alignment Scores:
Pred. No.: 0.876 Length: 4290
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 17 Gaps: 0

US-09-397-967-16 (1-1099) x AAT18394 (1-4290)
QY 941 ValHisArgAspLeuAlaAlaArgAsnIleuVal 952
DB 2618 GTCCACCGAGACCTGCTGCTCGCAACATCTAGTC 2653
RESULT 100
AAT51235
ID AAT51235 standard; cDNA to mRNA; 4290 BP.
XX
XX AAT51235;
AC
XX
XX 21-MAR-1997 (first entry)
XX
XX Receptor-type tyrosine kinase coding sequence.
XX
XX Receptor-type tyrosine kinase; LPM medium; lysand; autophosphorylation;
XX insect cell; animal cell; growth promoter; undifferentiated cell;
XX insulin; interleukin; fibroblast growth factor; hepatocyte growth factor;
XX nerve growth factor; interferon-gamma; tumour necrosis factor; inducer;
XX ss.
XX
XX Homo sapiens.
XX
XX
XX Key Location/Qualifiers
XX CDS 410..3373
XX /*tag= a
XX /*product= receptor-type tyrosine kinase

```

```

FT sig-peptide 410..454
FT /*tag= b
FT mat-peptide 455..3370
FT /*tag= c
PN JP08280385-A.
PD 29-OCT-1996.
XX
XX 13-APR-1995; 95JP-0087878.
XX
XX 13-APR-1995; 95JP-0087878.
XX
XX (ASAH ) ASAH KASEI KOGYO KK.
XX
XX WPI: 1997-014848/02.
XX P-PSDB: AAW11304.
XX
XX LPM cell culture medium - for culturing cells which produce
XX receptor-type tyrosine kinase lysand
XX
XX Claim 2; Page 34-38; 45pp; Japanese.
XX
XX This sequence represents the coding sequence for human receptor-type
XX tyrosine kinase. The protein encoded by this sequence is included in the
XX LPM medium of the invention. The medium is used to culture cells capable
XX of producing a lysand for receptor-type tyrosine kinase. The lysand
XX serves as a physiologically active substance to induce enzyme activity
XX and autophosphorylation of tyrosine kinase, which is involved in the
XX development and differentiation of animal and insect cells. The lysand
XX can also be used as in a pharmaceutical preparation for promoting growth
XX of undifferentiated cells. The medium may also include an inducer, such
XX as insulin, an interleukin, a fibroblast growth factor, or
XX interferon-gamma.
XX
SQ Sequence 4290 BP; 824 A; 1310 C; 1342 G; 809 T; 5 other;

Alignment Scores:
Pred. No.: 0.876 Length: 4290
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 1
DB: 18 Gaps: 0

US-09-397-967-16 (1-1099) x AAT51235 (1-4290)
QY 941 ValHisArgAspLeuAlaAlaArgAsnIleuVal 952
DB 2618 GTCCACCGAGACCTGCTGCTCGCAACATCTAGTC 2653

```

Search completed: April 28, 2003, 19:08:25
 Job time : 695 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.5
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: April 28, 2003, 21:08:40 ; Search time 267 Seconds

(without alignments)
4478.852 Million cell updates/sec

Title: US-09-397-967-16

Perfect score: 1099

Sequence: 1 MAPSETPLIPORSCSLSS.....RPAATATSPQDPLMRGRPG 1099

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 709820 seqs, 544064369 residues

Word size: 1

Total number of hits satisfying chosen parameters: 1417911

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 150 summaries

Command line parameters:

-MODEL=frame+...p2n.model -DEV=xlp
-O=/cg2_1/USPTO_spool/US09397967/runat_26042003_182929_11324/app.query.fasta_1.1287
-DB=Published Applications NA -OFMT=fastap -SUFFIX=oligo.rnpb -MINMATCH=0.1
-LOOPCL=0 -IOEXT=0 -UNITS=bits -STAR=1 -END=1 -MATRIX=oligo
-TRANS=human40.cdd -LIST=150 -DOCALIGN=200 -THR.SCORE=quality -THR.MIN=1
-ALIGN=100 -MODE=LOCAL -OUTEXT=ptc -NCRM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09397967@CON_1.1.175.@runat_26042003_182929_11324
-NCPU=6 -ICPU=3 -NO_XLPXY -NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT
-DISPLOCK=100 -LONGLOG -DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=60
-XGAPEXT=60 -Fgapop=6 -Fgapext=7 -YGAPOP=60 -YGAPEXT=60 -DELop=6 -DELExt=7

Database :

Published Applications_NA.*
1: /cg2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cg2_6/ptodata/1/pubpna/PCIT_NEW_PUB.seq.*
3: /cg2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cg2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cg2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cg2_6/ptodata/1/pubpna/PCITUS_PUBCOMB.seq.*
7: /cg2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cg2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cg2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
10: /cg2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
11: /cg2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
12: /cg2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
13: /cg2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
14: /cg2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	5.3	203	10	US-09-728-446-1413
2	4.6	439	9	US-09-918-995-11571
3	2.5	438	9	US-09-918-995-10879
4	1.6	364	10	US-09-771-161A-17

5	1.1	151	10	US-09-982-610-11	Sequence 11, Appl
6	1.1	294	10	US-09-864-761-21086	Sequence 21086, A
7	1.1	367	10	US-09-864-761-4335	Sequence 4335, Ap
8	1.1	372	10	US-09-878-574-1581	Sequence 1581, Ap
9	1.1	477	9	US-09-918-995-2066	Sequence 2066, Ap
10	1.1	917	9	US-09-991-936-1577	Sequence 1577, Ap
11	1.1	2000	9	US-09-977-260-1	Sequence 1, Appl1
12	1.1	2000	10	US-09-977-261-1	Sequence 1, Appl1
13	1.1	3107	10	US-09-977-269-1	Sequence 1, Appl1
14	1.1	3348	10	US-09-964-456-1128	Sequence 1128, Ap
15	1.1	3370	10	US-09-982-610-34	Sequence 34, Appl
16	1.1	3370	10	US-09-967-768A-144	Sequence 144, Appl
17	1.1	3393	10	US-09-922-138-3	Sequence 3, Appl1
18	1.1	3583	9	US-10-245-103-103	Sequence 103, Appl
19	1.1	3583	9	US-10-245-103-103	Sequence 103, Appl
20	1.1	3583	9	US-10-245-103-103	Sequence 103, Appl
21	1.1	3583	9	US-10-245-103-103	Sequence 103, Appl
22	1.1	3583	9	US-10-245-103-103	Sequence 103, Appl
23	1.1	3583	9	US-10-245-103-103	Sequence 103, Appl
24	1.1	3583	9	US-10-245-103-103	Sequence 103, Appl
25	1.1	3583	9	US-10-245-103-103	Sequence 103, Appl
26	1.1	3583	9	US-10-245-103-103	Sequence 103, Appl
27	1.1	3583	9	US-10-245-103-103	Sequence 103, Appl
28	1.1	3583	9	US-10-245-103-103	Sequence 103, Appl
29	1.1	3583	9	US-10-245-103-103	Sequence 103, Appl
30	1.1	3583	9	US-10-245-103-103	Sequence 103, Appl
31	1.1	3583	9	US-10-245-103-103	Sequence 103, Appl
32	1.1	3583	9	US-10-245-103-103	Sequence 103, Appl
33	1.1	3583	9	US-10-245-103-103	Sequence 103, Appl
34	1.1	3633	10	US-09-922-138-1	Sequence 1, Appl1
35	1.1	3945	10	US-09-921-771-4	Sequence 4, Appl1
36	1.1	3969	10	US-09-982-610-23	Sequence 23, Appl1
37	1.1	4698	10	US-09-895-652-1	Sequence 1, Appl1
38	1.1	5816	10	US-09-205-658-11	Sequence 11, Appl
39	1.1	5816	10	US-09-844-353A-11	Sequence 11, Appl
40	1.0	147	10	US-09-982-610-7	Sequence 7, Appl1
41	1.0	3018	9	US-09-938-842A-669	Sequence 669, Appl
42	1.0	4165	8	US-08-578-684-1	Sequence 1, Appl1
43	0.9	510	9	US-09-796-692-8268	Sequence 8268, Ap
44	0.9	536	9	US-10-101-464A-355	Sequence 355, Ap
45	0.9	539	9	US-09-796-692-8255	Sequence 8255, Ap
46	0.9	556	9	US-09-796-692-8283	Sequence 7883, Ap
47	0.9	564	9	US-09-796-692-9025	Sequence 9025, Ap
48	0.9	589	9	US-09-796-692-8630	Sequence 8630, Ap
49	0.9	597	9	US-09-796-692-7976	Sequence 7976, Ap
50	0.9	601	9	US-09-796-692-8273	Sequence 8273, Ap
51	0.9	605	9	US-09-796-692-8047	Sequence 8047, Ap
52	0.9	605	9	US-09-796-692-8927	Sequence 8927, Ap
53	0.9	608	9	US-09-796-692-8085	Sequence 8085, Ap
54	0.9	610	9	US-09-796-692-7538	Sequence 7538, Ap
55	0.9	645	9	US-10-101-464A-368	Sequence 368, Ap
56	0.9	869	9	US-10-101-464A-218	Sequence 218, Ap
57	0.9	1021	9	US-10-101-464A-343	Sequence 343, Ap
58	0.9	1128	9	US-10-101-464A-872	Sequence 872, Ap
59	0.9	2040	9	US-09-938-842A-1497	Sequence 1497, Ap
60	0.9	2052	9	US-09-938-842A-1480	Sequence 1480, Ap
61	0.9	2836	10	US-09-924-859A-4	Sequence 4, Appl1
62	0.9	3120	10	US-09-982-610-22	Sequence 22, Appl
63	0.9	3270	10	US-09-866-510-1	Sequence 1, Appl1
64	0.9	3270	10	US-09-866-510-3	Sequence 3, Appl1
65	0.9	3270	10	US-09-866-510-9	Sequence 9, Appl1
66	0.9	3270	10	US-09-866-510-5	Sequence 5, Appl1
67	0.9	3321	10	US-09-866-510-13	Sequence 13, Appl
68	0.9	3321	10	US-09-866-510-15	Sequence 15, Appl
69	0.9	3321	10	US-09-866-510-17	Sequence 17, Appl
70	0.9	3321	10	US-09-866-510-21	Sequence 21, Appl
71	0.9	3433	10	US-09-919-408-1	Sequence 1, Appl1
72	0.9	3433	10	US-09-872-136-1	Sequence 1, Appl1
73	0.9	3453	10	US-09-919-408-3	Sequence 3, Appl1
74	0.9	3501	10	US-09-872-136-3	Sequence 3, Appl1
75	0.9	4054	9	US-09-955-163-35	Sequence 35, Appl
76	0.9	4054	9	US-10-022-939-1	Sequence 1, Appl1
77	0.9	4071	9		

78	10	0	9	4071	9	US-10-100-4054-1	Sequence 1, Appl1
79	10	0	9	4111	9	US-09-375-248-1	Sequence 1, Appl1
80	10	0	9	4425	10	US-09-982-610-31	Sequence 31, Appl1
81	10	0	9	4465	9	US-09-955-363-11	Sequence 1, Appl1
82	10	0	9	4989	9	US-09-870-759-119	Sequence 119, Appl1
83	10	0	9	5084	10	US-09-954-531-1383	Sequence 1383, Ap
84	10	0	9	5084	9	US-09-967-768A-277	Sequence 277, App
85	10	0	9	5406	10	US-09-919-408-5	Sequence 5, Appl1
86	10	0	9	5406	10	US-09-872-136-5	Sequence 5, Appl1
87	10	0	9	5470	10	US-09-766-678-1	Sequence 1, Appl1
88	10	0	9	6378	10	US-09-919-497-40	Sequence 40, Appl1
89	10	0	9	6412	10	US-09-769-987-1	Sequence 1, Appl1
90	10	0	9	6827	10	US-09-982-610-17	Sequence 17, Appl1
91	10	0	9	9108	10	US-09-982-610-45	Sequence 45, Appl1
92	9	0	8	3352	10	US-09-960-352-110875	Sequence 108775, Ap
93	9	0	8	3352	10	US-09-960-352-1117	Sequence 1177, Ap
94	9	0	8	363	9	US-10-101-464A-332	Sequence 332, App
95	9	0	8	366	10	US-09-960-352-10878	Sequence 10878, App
96	9	0	8	391	10	US-09-960-352-12603	Sequence 2603, Ap
97	9	0	8	392	10	US-09-960-352-4293	Sequence 4293, Ap
98	9	0	8	401	10	US-09-960-352-13765	Sequence 13765, Ap
99	9	0	8	402	10	US-09-960-352-14442	Sequence 14442, A
100	9	0	8	415	9	US-09-918-995-16302	Sequence 16302, A
101	9	0	8	415	10	US-09-960-352-10423	Sequence 10423, A
102	9	0	8	422	10	US-09-960-352-6596	Sequence 6596, Ap
103	9	0	8	434	10	US-09-960-352-5227	Sequence 5227, Ap
104	9	0	8	462	10	US-09-864-761-20164	Sequence 20164, A
105	9	0	8	463	9	US-10-102-806-228	Sequence 228, App
106	9	0	8	463	9	US-09-918-995-23959	Sequence 23959, A
107	9	0	8	480	9	US-10-060-036-4148	Sequence 4148, Ap
108	9	0	8	482	9	US-09-796-692-6991	Sequence 6991, Ap
109	9	0	8	486	9	US-10-101-464A-378	Sequence 378, App
110	9	0	8	513	9	US-09-918-995-14461	Sequence 14461, A
111	9	0	8	753	10	US-09-910-943-636	Sequence 636, App
112	9	0	8	903	10	US-09-815-242-4130	Sequence 4130, Ap
113	9	0	8	1620	9	US-10-118-984-3	Sequence 3, Appl1
114	9	0	8	1620	10	US-09-728-721-3	Sequence 3, Appl1
115	9	0	8	1620	12	US-10-105-931-3	Sequence 3, Appl1
116	9	0	8	1755	9	US-09-930-125-6	Sequence 6, Appl1
117	9	0	8	1767	9	US-09-930-125-4	Sequence 4, Appl1
118	9	0	8	1773	9	US-09-930-125-7	Sequence 7, Appl1
119	9	0	8	1806	9	US-09-930-125-5	Sequence 5, Appl1
120	9	0	8	1931	10	US-10-118-984-1	Sequence 1, Appl1
121	9	0	8	1931	10	US-09-748-537-2	Sequence 2, Appl1
122	9	0	8	1931	10	US-09-728-721-1	Sequence 1, Appl1
123	9	0	8	1931	12	US-10-133-780-2	Sequence 2, Appl1
124	9	0	8	1931	12	US-10-105-931-1	Sequence 1, Appl1
125	9	0	8	2187	10	US-09-728-952-82	Sequence 82, Appl1
126	9	0	8	2187	9	US-09-954-531-188	Sequence 188, App
127	9	0	8	2338	10	US-09-925-302-90	Sequence 90, Appl1
128	9	0	8	2437	9	US-09-158-722-3	Sequence 3, Appl1
129	9	0	8	2454	10	US-09-964-824A-120	Sequence 120, App
130	9	0	8	2627</			

[illegible]


```

      : TELEX: 910/371-7168
      : INFORMATION FOR SEQ ID NO: 11:
      : SEQUENCE CHARACTERISTICS:
      :     LENGTH: 151 base pairs
      :     TYPE: Nucleic Acid
      :     STRANDEDNESS: Single
      :     TOPOLOGY: Linear
      : SEQUENCE DESCRIPTION: SEQ ID NO: 11:
      : US-09-982-610-11

Alignment Scores:
Pred. No.:          0.0027          Length:          151
Score:             12.00           Matches:         12
Percent Similarity: 100.00%       Conservative:    0
Best local Similarity: 100.00%   Mismatches:     0
Query Match:       1.09%         Indels:          0
DB:                10           Gaps:            0

US-09-397-967-16 (1-1099) x US-09-982-610-11 (1-151)
Oy      941 ValHisArgAspLeuAlaIalaArgasnileLeuVal 952
        |||||||
Db       1 GTGCACAGGGATCTCGCGGCTCGGAACATCTCTC 36

RESULT 6
US-09-864-761-21086
: Sequence 21086, Application US/09864761
: Patent No. US20020048763A1
: GENERAL INFORMATION:
: APPLICANT: Penn, Sharon G.
: APPLICANT: Rank, David R.
: APPLICANT: Hanzel, David K.
: APPLICANT: Chen, Wensheng
: TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
: FILE REFERENCE: Aecomica-x-1
: CURRENT APPLICATION NUMBER: US/09/864,761
: PRIOR FILING DATE: 2001-05-23
: PRIOR APPLICATION NUMBER: US 60/180,312
: PRIOR FILING DATE: 2000-02-04
: PRIOR APPLICATION NUMBER: US 60/207,456
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: US 09/632,366
: PRIOR FILING DATE: 2000-08-03
: PRIOR APPLICATION NUMBER: GB 24263.6
: PRIOR FILING DATE: 2000-10-04
: PRIOR APPLICATION NUMBER: US 60/236,359
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: PCT/US01/00666
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00667
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00664
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00669
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00665
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00668
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00663
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00662
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00661
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00670
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: US 60/234,687
: PRIOR FILING DATE: 2000-09-21
: PRIOR APPLICATION NUMBER: US 09/608,408
: PRIOR FILING DATE: 2000-06-30
: PRIOR APPLICATION NUMBER: US 09/774,203

```

```

PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1
SEQ ID NO 21086
LENGTH: 294
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
  OTHER INFORMATION: MAP TO AC010899.2
  OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.6
  OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.7
  OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 4.5
  OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.3
  OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 3.8
  OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 4.3
  OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.8
  OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.2
  OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 3.2
  OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 4.4
  OTHER INFORMATION: SWISSPROT HIT: Q91694, EVALUE 2.00e-35
  OTHER INFORMATION: NT HIT: g14758279, EVALUE 1.00e-115
  OTHER INFORMATION: EST_HUMAN HIT: AW65297.1, EVALUE 2.00e-26
US-09-864-761-21086

Alignment Scores:
Pred. No.: 0.0049 Length: 294
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
Db: Gaps: 0

US-09-397-967-16 (1-1099) x US-09-864-761-21086 (1-294)
Qy 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
Db 91 GTGCATCTGATCTGCGCCGACGGACATCTCTGGTG 126

RESULT 7
US-09-864-761-4335
Sequence 4335, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
  APPLICANT: Penn, Sharron G.
  APPLICANT: Hanzel, David R.
  APPLICANT: Chen, Wensheng
  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
  FILE REFERENCE: Acomica-X-1
  CURRENT APPLICATION NUMBER: US/09/864,761
  CURRENT FILING DATE: 2001-05-23
  PRIOR APPLICATION NUMBER: US 60/180,312
  PRIOR FILING DATE: 2000-02-04
  PRIOR APPLICATION NUMBER: US 60/207,456
  PRIOR FILING DATE: 2000-05-26
  PRIOR APPLICATION NUMBER: US 09/632,366
  PRIOR FILING DATE: 2000-08-03
  PRIOR APPLICATION NUMBER: GB 24263,6
  PRIOR FILING DATE: 2000-10-04
  PRIOR APPLICATION NUMBER: US 60/236,359
  PRIOR FILING DATE: 2000-09-27
  PRIOR APPLICATION NUMBER: PCT/US01/00666
  PRIOR FILING DATE: 2001-01-30
  PRIOR APPLICATION NUMBER: PCT/US01/00667
  PRIOR FILING DATE: 2001-01-30
  PRIOR APPLICATION NUMBER: PCT/US01/00664
  PRIOR FILING DATE: 2001-01-30
  PRIOR APPLICATION NUMBER: PCT/US01/00669
  PRIOR FILING DATE: 2001-01-30
  PRIOR APPLICATION NUMBER: PCT/US01/00665
  PRIOR FILING DATE: 2001-01-30
  PRIOR APPLICATION NUMBER: PCT/US01/00668

```



```

; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 4335
; LENGTH: 367
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC010899.2
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.6
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.7
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 4.5
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.3
; OTHER INFORMATION: EXPRESSED IN HEL100, SIGNAL = 3.8
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 4.3
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.8
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.2
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 3.2
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 4.4
; US-09-864-761-4335

Alignment Scores:
Pred. No.: 0.00597 Length: 367
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-864-761-4335 (1-367)

QY 941 VALHISARGSPLEUALAALARGASNIIEUVAL 952
DB 289 GTGCATCGTGCATCTGCGCCGACGAAACATCTGTGTC 324

RESULT 8
; Sequence 1581, Application US/09878574
; Patent No. US20020110548A1
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: La Rosa, Thomas J.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(15401)B
; CURRENT APPLICATION NUMBER: US/09/878,574
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/333,535
; PRIOR FILING DATE: 1999-06-14
; NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 1581
; LENGTH: 372
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: L1B3028-036-01-B1-B10
; US-09-878-574-1581

Alignment Scores:
```

```

Pred. No.: 0.00605 Length: 372
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-878-574-1581 (1-372)

QY 1003 SERSAPVALTRPSRPHGLYVALLEUITYRGU 1014
DB 329 AGTGATGTGTGAGTGTGTTGGGTGGTGTATATGAA 364

RESULT 9
; Sequence 2066, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2066
; LENGTH: 477
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)-(477)
; OTHER INFORMATION: n - A,T,C or G
; US-09-918-995-2066

Alignment Scores:
Pred. No.: 0.00755 Length: 477
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 9 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-918-995-2066 (1-477)

QY 941 VALHISARGSPLEUALAALARGASNIIEUVAL 952
DB 200 GTGCACCGCGACCTGCTGCTGCAACATCTGTGTC 235

RESULT 10
; Sequence 1577, Application US/09991936
; Publication No. US20030073627A1
; GENERAL INFORMATION:
; APPLICANT: Brandt, Kevin S.
; APPLICANT: Gaines, Patrick J.
; APPLICANT: Stinchcomb, Dan T.
; APPLICANT: Miszewski, Nancy
; TITLE OF INVENTION: FLEA HEAD, NERVE CORD, HINDGUT AND MALPIGHIAN TUBULE
; FILE REFERENCE: FC-6-C1
; CURRENT APPLICATION NUMBER: US/09/991,936
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US/09/543,668
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 60/128,704
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 1959
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1577
```

```
; LENGTH: 917
; TYPE: DNA
; ORGANISM: Ctenocephalides felis
US-09-991-936-1577

Alignment Scores:
Pred. No.: 0.0136      Length: 917
Score: 12.00           Matches: 12
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match: 1.09%             Indels: 0
DB: 9                    Gaps: 0

US-09-397-967-16 (1-1099) x US-09-991-936-1577 (1-917)

OY 941 ValHisArgAspLeuAlaAlaArgAsnIleuVal 952
|||||
Db 169 GTCCATCGAGATTGGCTGCCAGGAAATATCTGTG 204

RESULT 11
US-09-977-260-1
; Sequence 1, Application US/09977260
; Publication No. US20020192790A1
; GENERAL INFORMATION:
; APPLICANT: ULRICH, AXEL
; APPLICANT: GISHIZKY, MIKHAIL
; TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN TYROSINE KINASES
; FILE REFERENCE: 038602/1260
; CURRENT APPLICATION NUMBER: US/09/977,260
; PRIOR FILING DATE: 2001-10-16
; PRIOR APPLICATION NUMBER: 08/232,545
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Unknown Organism
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (258)..(1778)
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Megakaryocyte
; OTHER INFORMATION: kinase 1
US-09-977-260-1

Alignment Scores:
Pred. No.: 0.0272      Length: 2000
Score: 12.00           Matches: 12
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match: 1.09%             Indels: 0
DB: 9                    Gaps: 0

US-09-397-967-16 (1-1099) x US-09-977-260-1 (1-2000)

OY 941 ValHisArgAspLeuAlaAlaArgAsnIleuVal 952
|||||
Db 1302 GTGCACCGCGACCTGGCGCCGCCGCAACATCTGTG 1337

RESULT 12
US-09-977-261-1
; Sequence 1, Application US/09977261
; Publication No. US20030054527A1
; GENERAL INFORMATION:
; APPLICANT: ULRICH, AXEL
; APPLICANT: GISHIZKY, MIKHAIL
; TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN TYROSINE KINASES
; FILE REFERENCE: 038602/1259
; CURRENT APPLICATION NUMBER: US/09/977,261
; CURRENT FILING DATE: 2001-10-16
```

```
; PRIOR APPLICATION NUMBER: 08/232,545
; PRIOR FILING DATE: 1994-04-22
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Unknown Organism
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (258)..(1778)
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Megakaryocyte
; OTHER INFORMATION: kinase 1
US-09-977-261-1

Alignment Scores:
Pred. No.: 0.0272      Length: 2000
Score: 12.00           Matches: 12
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match: 1.09%             Indels: 0
DB: 9                    Gaps: 0

US-09-397-967-16 (1-1099) x US-09-977-261-1 (1-2000)

OY 941 ValHisArgAspLeuAlaAlaArgAsnIleuVal 952
|||||
Db 1302 GTGCACCGCGACCTGGCGCCGCCGCAACATCTGTG 1337

RESULT 13
US-09-977-269-1
; Sequence 1, Application US/09977269
; Patent No. US20020082037A1
; GENERAL INFORMATION:
; APPLICANT: ULRICH, AXEL
; APPLICANT: GISHIZKY, MIKHAIL
; TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN TYROSINE KINASES
; FILE REFERENCE: 038602/1260
; CURRENT APPLICATION NUMBER: US/09/977,269
; PRIOR FILING DATE: 2001-10-16
; PRIOR APPLICATION NUMBER: 08/232,545
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Unknown Organism
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (258)..(1778)
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Megakaryocyte
; OTHER INFORMATION: kinase 1
US-09-977-269-1

Alignment Scores:
Pred. No.: 0.0272      Length: 2000
Score: 12.00           Matches: 12
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match: 1.09%             Indels: 0
DB: 10                   Gaps: 0

US-09-397-967-16 (1-1099) x US-09-977-269-1 (1-2000)

OY 941 ValHisArgAspLeuAlaAlaArgAsnIleuVal 952
|||||
Db 1302 GTGCACCGCGACCTGGCGCCGCCGCAACATCTGTG 1337

RESULT 14
```

US-09-954-456-1128
; Sequence 1128, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cand
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1128
; LENGTH: 3107
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-456-1128

Alignment Scores:
Pred. No.: 0.0404 Length: 3107
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-954-456-1128 (1-3107)

QY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
DB 2260 GTGCATCGTGATCTGCGCCGACGAGACATCTGCTG 2295
|||||

RESULT 15
US-09-982-610-34
; Sequence 34, Application US/09982610
; Patent No. US20020146420A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; Bennett, Brian D.
; Goeddel, David
; Lee, James M.
; Matthews, William
; Tsai, Siao Ping
; Wood, William I.
; TITLE OF INVENTION: PROTEIN TYROSINE KINASE AGONIST ANTIBODIES
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Winpatin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/982,610
FILING DATE: 17-Oct-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/446,648
FILING DATE: 1996-MAY-23
APPLICATION NUMBER: 08/222616
FILING DATE: 04-APR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M.
REGISTRATION NUMBER: 40,378
REFERENCE/DOCKET NUMBER: P0821P3PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1994
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 3348 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Single
TOPOLOGY: Linear
SEQUENCE DESCRIPTION: SEQ ID NO: 34:
US-09-982-610-34

Alignment Scores:
Pred. No.: 0.0432 Length: 3348
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-982-610-34 (1-3348)

QY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
DB 2227 GTGCATCGTGATCTGCGCCGACGAGACATCTGCTG 2262
|||||

RESULT 16
US-09-967-768A-144
; Sequence 144, Application US/09967768A
; Patent No. US20020150877A1
; GENERAL INFORMATION:
; APPLICANT: Augustus, Meena
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Sign
; FILE REFERENCE: 689290-72
; CURRENT APPLICATION NUMBER: US/09/967,768A
; FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US/60/236,109
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,034
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,111
; PRIOR FILING DATE: 2000-09-28
; NUMBER OF SEQ ID NOS: 325
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 144
; LENGTH: 3370
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-967-768A-144

Alignment Scores:
Pred. No.: 0.0435 Length: 3370
Score: 12.00 Matches: 12

Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-967-768A-144 (1-3370)

QY 941 VALHISARGSPLEUALAALARGASNIIEUVAL 952
DB 2353 GTCCACGGGAGCTGCTGCCAGAAACATCTTGTC 2388

RESULT 17
US-09-922-138-3
; Sequence 3, Application US/09922138
; Patent No. US20020061574A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Meyers, Rachel
; APPLICANT: Silos-Santiago, Immaculada
; TITLE OF INVENTION: 1658, 14223, AND 16002, NOVEL HUMAN
; TITLE OF INVENTION: KINASES AND USES THEREFOR
; FILE REFERENCE: 38155-20030.00
; CURRENT APPLICATION NUMBER: US/09/922.138
; CURRENT FILING DATE: 2001-11-28
; PRIOR APPLICATION NUMBER: US 60/729, 299
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 3393
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-922-138-3

Alignment Scores:
Pred. No.: 0.0437 Length: 3393
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-922-138-3 (1-3393)

QY 941 VALHISARGSPLEUALAALARGASNIIEUVAL 952
DB 2665 GTTCATCGAGACCTAGCGGCTCGGAATATACTGTC 2700

RESULT 18
US-10-245-103-103
; Sequence 103, Application US/10245103
; Publication No. US20030068778A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Eaton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Phillippe
; APPLICANT: Watanabe, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3630R1C112
; CURRENT APPLICATION NUMBER: US/10/245.103
; CURRENT FILING DATE: 2002-09-17
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114

; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 103
; LENGTH: 3583
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-245-103-103

Alignment Scores:
Pred. No.: 0.0459 Length: 3583
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 9 Gaps: 0

US-09-397-967-16 (1-1099) x US-10-245-103-103 (1-3583)

QY 941 VALHISARGSPLEUALAALARGASNIIEUVAL 952
DB 2633 GTTCATCGAGACCTAGCGGCTCGGAATATACTGTC 2668

RESULT 19
US-10-245-107-103
; Sequence 103, Application US/10245107
; Publication No. US20030068779A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Eaton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Phillippe
; APPLICANT: Watanabe, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3630R1C71
; CURRENT APPLICATION NUMBER: US/10/245.107
; CURRENT FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22

```

: PRIOR APPLICATION NUMBER: 60/087607
: PRIOR FILING DATE: 1998-06-02
: PRIOR APPLICATION NUMBER: 60/089801
: PRIOR FILING DATE: 1998-06-18
: PRIOR APPLICATION NUMBER: 60/090557
: PRIOR FILING DATE: 1998-06-24
: PRIOR APPLICATION NUMBER: 60/090689
: PRIOR FILING DATE: 1998-06-25
: Remaining Prior Application data removed - See File Wrapper or PALM.
: SEQ ID NO 103
: LENGTH: 3583
: TYPE: DNA
: ORGANISM: Homo Sapien
US-10-245-107-103

Alignment Scores:
Pred. No.: 0.0459      Length: 3583
Score: 12.00           Matches: 12
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match: 1.09%              Indels: 0
DB: 9                    Gaps: 0

US-09-397-967-16 (1-1099) x US-10-245-107-103 (1-3583)

QY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
Db 2633 GTTCATCGAGACCTAGCGGCTCGGAATATATCTGCTC 2668

RESULT 20
US-10-245-143-103
: Sequence 103, Application US/10245143
: Publication No. US20030068780A1
: GENERAL INFORMATION:
: APPLICANT: Baker, Kevin
: APPLICANT: Baton, Dan
: APPLICANT: Filvaroff, Ellen
: APPLICANT: Goddard, Audrey
: APPLICANT: Grimaldi, J. Christopher
: APPLICANT: Gurney, Austin
: APPLICANT: Smith, Victoria
: APPLICANT: Stephan, Jean-Philippe
: APPLICANT: Watande, Colin
: APPLICANT: Wood, William
: APPLICANT: Zhang, Zemin
: APPLICANT: Fong, Sherman
: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
: TITLE OF INVENTION: ACIDS ENCODING THE SAME
: FILE REFERENCE: P3630R1C90
: CURRENT APPLICATION NUMBER: US/10/245,143
: CURRENT FILING DATE: 2002-09-16
: PRIOR APPLICATION NUMBER: 10/197942
: PRIOR FILING DATE: 2002-07-18
: PRIOR APPLICATION NUMBER: 60/059114
: PRIOR FILING DATE: 1997-09-17
: PRIOR APPLICATION NUMBER: 60/063046
: PRIOR FILING DATE: 1997-10-24
: PRIOR APPLICATION NUMBER: 60/065027
: PRIOR FILING DATE: 1997-11-10
: PRIOR APPLICATION NUMBER: 60/079689
: PRIOR FILING DATE: 1998-03-27
: PRIOR APPLICATION NUMBER: 60/086478
: PRIOR FILING DATE: 1998-05-22
: PRIOR APPLICATION NUMBER: 60/087607
: PRIOR FILING DATE: 1998-06-02
: PRIOR APPLICATION NUMBER: 60/089801
: PRIOR FILING DATE: 1998-06-18
: PRIOR APPLICATION NUMBER: 60/090557
: PRIOR FILING DATE: 1998-06-24
: PRIOR APPLICATION NUMBER: 60/090689
: PRIOR FILING DATE: 1998-06-25
: Remaining Prior Application data removed - See File Wrapper or PALM.
: Remaining Prior Application data removed - See File Wrapper or PALM.
```

```

: NUMBER OF SEQ ID NOS: 116
: SEQ ID NO 103
: LENGTH: 3583
: TYPE: DNA
: ORGANISM: Homo Sapien
US-10-245-143-103

Alignment Scores:
Pred. No.: 0.0459      Length: 3583
Score: 12.00           Matches: 12
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match: 1.09%              Indels: 0
DB: 9                    Gaps: 0

US-09-397-967-16 (1-1099) x US-10-245-143-103 (1-3583)

QY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
Db 2633 GTTCATCGAGACCTAGCGGCTCGGAATATATCTGCTC 2668

RESULT 21
US-10-245-771-103
: Sequence 103, Application US/10245771
: Publication No. US20030068781A1
: GENERAL INFORMATION:
: APPLICANT: Baker, Kevin
: APPLICANT: Baton, Dan
: APPLICANT: Filvaroff, Ellen
: APPLICANT: Goddard, Audrey
: APPLICANT: Grimaldi, J. Christopher
: APPLICANT: Gurney, Austin
: APPLICANT: Smith, Victoria
: APPLICANT: Stephan, Jean-Philippe
: APPLICANT: Watande, Colin
: APPLICANT: Wood, William
: APPLICANT: Zhang, Zemin
: APPLICANT: Fong, Sherman
: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
: TITLE OF INVENTION: ACIDS ENCODING THE SAME
: FILE REFERENCE: P3630R1C98
: CURRENT APPLICATION NUMBER: US/10/245,771
: CURRENT FILING DATE: 2002-09-16
: PRIOR APPLICATION NUMBER: 10/197942
: PRIOR FILING DATE: 2002-07-18
: PRIOR APPLICATION NUMBER: 60/059114
: PRIOR FILING DATE: 1997-09-17
: PRIOR APPLICATION NUMBER: 60/063046
: PRIOR FILING DATE: 1997-10-24
: PRIOR APPLICATION NUMBER: 60/065027
: PRIOR FILING DATE: 1997-11-10
: PRIOR APPLICATION NUMBER: 60/079689
: PRIOR FILING DATE: 1998-03-27
: PRIOR APPLICATION NUMBER: 60/086478
: PRIOR FILING DATE: 1998-05-22
: PRIOR APPLICATION NUMBER: 60/087607
: PRIOR FILING DATE: 1998-06-02
: PRIOR APPLICATION NUMBER: 60/089801
: PRIOR FILING DATE: 1998-06-18
: PRIOR APPLICATION NUMBER: 60/090557
: PRIOR FILING DATE: 1998-06-24
: PRIOR APPLICATION NUMBER: 60/090689
: PRIOR FILING DATE: 1998-06-25
: Remaining Prior Application data removed - See File Wrapper or PALM.
: NUMBER OF SEQ ID NOS: 116
: SEQ ID NO 103
: LENGTH: 3583
: TYPE: DNA
: ORGANISM: Homo Sapien
US-10-245-771-103

Alignment Scores:
Pred. No.: 0.0459      Length: 3583
```

Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 9 Gaps: 0

US-09-397-967-16 (1-1099) x US-10-245-771-103 (1-3583)

OY 941 VALHISARGASPLEUALAALARGASNILEUVAL 952

DB 2633 GTTCATCGAGACCTAGCGGCTCGGAATATACTGCTC 2668

RESULT 22

US-10-245-851-103

; Sequence 103, Application US/10245851
; Publication No. US20030068782A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin

; APPLICANT: Eaton, Dan

; APPLICANT: Filvaroff, Ellen

; APPLICANT: Goddard, Audrey

; APPLICANT: Grimaldi, J. Christopher

; APPLICANT: Gurney, Austin

; APPLICANT: Smith, Victoria

; APPLICANT: Stephan, Jean-Phillippe

; APPLICANT: Matande, Colin

; APPLICANT: Wood, William

; APPLICANT: Zhang, Zemin

; APPLICANT: Fong, Sherman

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; FILE REFERENCE: P3630R1C93

; CURRENT FILING DATE: 2002-09-16

; PRIOR APPLICATION NUMBER: 60/065027

; PRIOR FILING DATE: 1997-11-10

; PRIOR APPLICATION NUMBER: 60/086478

; PRIOR FILING DATE: 1998-06-18

; PRIOR APPLICATION NUMBER: 60/090557

; PRIOR FILING DATE: 1998-06-24

; PRIOR APPLICATION NUMBER: 60/090689

; PRIOR FILING DATE: 1998-06-25

; Remaining Prior Application data removed - See File Wrapper or PALM.

; SEQ ID NO 103

; LENGTH: 3583

; TYPE: DNA

; ORGANISM: Homo Sapien

; US-10-245-851-103

Alignment Scores:

Pred. No.: 0.0459

Score: 12.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 1.09%

DB: 9

US-09-397-967-16 (1-1099) x US-10-245-851-103 (1-3583)

OY 941 VALHISARGASPLEUALAALARGASNILEUVAL 952

DB 2633 GTTCATCGAGACCTAGCGGCTCGGAATATACTGCTC 2668

RESULT 23

US-10-245-883-103

; Sequence 103, Application US/10245883
; Publication No. US20030068783A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin

; APPLICANT: Eaton, Dan

; APPLICANT: Filvaroff, Ellen

; APPLICANT: Goddard, Audrey

; APPLICANT: Grimaldi, J. Christopher

; APPLICANT: Gurney, Austin

; APPLICANT: Smith, Victoria

; APPLICANT: Stephan, Jean-Phillippe

; APPLICANT: Matande, Colin

; APPLICANT: Wood, William

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; FILE REFERENCE: P3630R1C70

; CURRENT FILING DATE: 2002-09-16

; PRIOR APPLICATION NUMBER: 60/059114

; PRIOR FILING DATE: 1997-09-17

; PRIOR APPLICATION NUMBER: 60/063046

; PRIOR FILING DATE: 1997-10-24

; PRIOR APPLICATION NUMBER: 60/065027

; PRIOR FILING DATE: 1997-11-10

; PRIOR APPLICATION NUMBER: 60/079689

; PRIOR FILING DATE: 1998-03-27

; PRIOR APPLICATION NUMBER: 60/086478

; PRIOR FILING DATE: 1998-05-22

; PRIOR APPLICATION NUMBER: 60/087607

; PRIOR FILING DATE: 1998-06-02

; PRIOR APPLICATION NUMBER: 60/089801

; PRIOR FILING DATE: 1998-06-18

; PRIOR APPLICATION NUMBER: 60/090557

; PRIOR FILING DATE: 1998-06-24

; PRIOR APPLICATION NUMBER: 60/090689

; PRIOR FILING DATE: 1998-06-25

; Remaining Prior Application data removed - See File Wrapper or PALM.

; SEQ ID NO 103

; LENGTH: 3583

; TYPE: DNA

; ORGANISM: Homo Sapien

; US-10-245-883-103

Alignment Scores:

Pred. No.: 0.0459

Score: 12.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 1.09%

DB: 9

US-09-397-967-16 (1-1099) x US-10-245-883-103 (1-3583)

OY 941 VALHISARGASPLEUALAALARGASNILEUVAL 952

DB 2633 GTTCATCGAGACCTAGCGGCTCGGAATATACTGCTC 2668

RESULT 24

US-10-237-535-103

; Sequence 103, Application US/10237535
; Publication No. US20030073188A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin

APPLICANT: Eaton, Dan
APPLICANT: Filvarcoff, Ellen
APPLICANT: Goddard, Audrey
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin
APPLICANT: Smith, Victoria
APPLICANT: Stephan, Jean-Phillippe
APPLICANT: Watanabe, Colin
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
APPLICANT: Fong, Sherman
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3630R1C3
CURRENT APPLICATION NUMBER: US/10/237,535
PRIOR FILING DATE: 2002-09-06
PRIOR APPLICATION NUMBER: 10/197942
PRIOR FILING DATE: 2002-07-18
PRIOR APPLICATION NUMBER: 60/059114
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/063046
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/065027
PRIOR FILING DATE: 1997-11-10
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/086478
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090689
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/091358
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091978
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/099803
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/106932
PRIOR FILING DATE: 1998-11-03
PRIOR APPLICATION NUMBER: 60/115554
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/119342
PRIOR FILING DATE: 1999-02-09
PRIOR APPLICATION NUMBER: 60/123957
PRIOR FILING DATE: 1999-03-12
PRIOR APPLICATION NUMBER: 60/123972
PRIOR FILING DATE: 1999-03-11
PRIOR APPLICATION NUMBER: 60/127372
PRIOR FILING DATE: 1999-04-01
PRIOR APPLICATION NUMBER: 60/131271
PRIOR FILING DATE: 1999-04-27
PRIOR APPLICATION NUMBER: 60/133459
PRIOR FILING DATE: 1999-05-11
PRIOR APPLICATION NUMBER: 60/135725
PRIOR FILING DATE: 1999-05-25
PRIOR APPLICATION NUMBER: 60/135729
PRIOR FILING DATE: 1999-05-25
PRIOR APPLICATION NUMBER: 60/135750
PRIOR FILING DATE: 1999-05-25
PRIOR APPLICATION NUMBER: 60/138385
PRIOR FILING DATE: 1999-06-09
PRIOR APPLICATION NUMBER: 60/140653
PRIOR FILING DATE: 1999-06-22
PRIOR APPLICATION NUMBER: 60/141037
PRIOR FILING DATE: 1999-06-23
PRIOR APPLICATION NUMBER: 60/144732
PRIOR FILING DATE: 1999-07-20
PRIOR APPLICATION NUMBER: 60/144758

PRIOR FILING DATE: 1999-07-20
PRIOR APPLICATION NUMBER: 60/144790
PRIOR FILING DATE: 1999-07-20
PRIOR APPLICATION NUMBER: 60/145228
PRIOR FILING DATE: 1999-07-20
PRIOR APPLICATION NUMBER: 60/145698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: 60/146222
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: 60/146843
PRIOR FILING DATE: 1999-08-03
PRIOR APPLICATION NUMBER: 60/148188
PRIOR FILING DATE: 1999-08-10
PRIOR APPLICATION NUMBER: 60/148513
PRIOR FILING DATE: 1999-08-12
PRIOR APPLICATION NUMBER: 60/149327
PRIOR FILING DATE: 1999-08-17
PRIOR APPLICATION NUMBER: 60/149395
PRIOR FILING DATE: 1999-08-17
PRIOR APPLICATION NUMBER: 60/150114
PRIOR FILING DATE: 1999-08-20
PRIOR APPLICATION NUMBER: 60/151700
PRIOR FILING DATE: 1999-08-31
PRIOR APPLICATION NUMBER: 60/151734
PRIOR FILING DATE: 1999-08-31
PRIOR APPLICATION NUMBER: 60/162506
PRIOR FILING DATE: 1999-10-29
PRIOR APPLICATION NUMBER: 60/170262
PRIOR FILING DATE: 1999-12-09
PRIOR APPLICATION NUMBER: 60/177118
PRIOR FILING DATE: 2000-01-20
PRIOR APPLICATION NUMBER: 60/179851
PRIOR FILING DATE: 2000-02-02
PRIOR APPLICATION NUMBER: 60/180921
PRIOR FILING DATE: 2000-02-08
PRIOR APPLICATION NUMBER: 60/187202
PRIOR FILING DATE: 2000-03-03
PRIOR APPLICATION NUMBER: 60/198587
PRIOR FILING DATE: 2000-04-18
PRIOR APPLICATION NUMBER: 60/199614
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 60/206330
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/206368
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/209832
PRIOR FILING DATE: 2000-06-05
PRIOR APPLICATION NUMBER: 60/218371
PRIOR FILING DATE: 2000-07-13
PRIOR APPLICATION NUMBER: 60/222695
PRIOR FILING DATE: 2000-08-02
PRIOR APPLICATION NUMBER: 60/229896
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/230621
PRIOR FILING DATE: 2000-09-05
PRIOR APPLICATION NUMBER: 60/232887
PRIOR FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: 60/235147
PRIOR FILING DATE: 2000-09-22
PRIOR APPLICATION NUMBER: 60/261878
PRIOR FILING DATE: 2001-01-12
PRIOR APPLICATION NUMBER: 60/261910
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: 60/261939
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: 60/262150
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: 60/264395
PRIOR FILING DATE: 2001-01-25
PRIOR APPLICATION NUMBER: 60/266421
PRIOR FILING DATE: 2001-02-02
PRIOR APPLICATION NUMBER: 60/267623
PRIOR FILING DATE: 2001-02-09

```

; PRIOR APPLICATION NUMBER: 60/274399
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 60/280982
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: 60/282129
; PRIOR FILING DATE: 2001-04-04
; PRIOR APPLICATION NUMBER: 60/282199
; PRIOR FILING DATE: 2001-04-04
; PRIOR APPLICATION NUMBER: 60/290589
; PRIOR FILING DATE: 2001-05-09
; PRIOR APPLICATION NUMBER: 09/180997
; PRIOR FILING DATE: 1998-11-19
; PRIOR APPLICATION NUMBER: 09/267213
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: 09/380137
; PRIOR FILING DATE: 1999-08-25
; PRIOR APPLICATION NUMBER: 09/380138
; PRIOR FILING DATE: 1999-08-25
; PRIOR APPLICATION NUMBER: 09/403297
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: 09/423741
; PRIOR FILING DATE: 1999-11-10
; PRIOR APPLICATION NUMBER: 09/709238
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 09/802706
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 09/872035
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: 09/918585
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 09/924419
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: 09/927796
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 09/929404
; PRIOR FILING DATE: 2001-08-13
; PRIOR APPLICATION NUMBER: 09/931836
; PRIOR FILING DATE: 2001-08-16
; PRIOR APPLICATION NUMBER: 09/941992
; PRIOR FILING DATE: 2001-08-28
; PRIOR APPLICATION NUMBER: 09/946374
; PRIOR FILING DATE: 2001-09-04
; PRIOR APPLICATION NUMBER: 10/001054
; PRIOR FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 10/081056
; PRIOR FILING DATE: 2002-02-20
; PRIOR APPLICATION NUMBER: 10/119480
; PRIOR FILING DATE: 2002-04-09

Alignment Scores:
Pred. No.: 0.0459 Length: 3583
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
Gaps: 0

US-09-397-967-16 (1-1099) x US-10-237-535-103 (1-3583)
OY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
Db 2633 GTTCATCGAGACCTAGCGCTCGGATATACTGCTC 2668

RESULT 25
US-10-238-183-103
; Sequence 103, Application US/10238183
; Publication No. US20030073189A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Eaton, Dan
; APPLICANT: Filvaroff, Ellen
```

```

; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe
; APPLICANT: Watande, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3630R1C11
; CURRENT FILING DATE: 2002-09-09
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/091358
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/099803
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/106932
; PRIOR FILING DATE: 1998-11-03
; PRIOR APPLICATION NUMBER: 60/115554
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/119342
; PRIOR FILING DATE: 1999-02-09
; PRIOR APPLICATION NUMBER: 60/123957
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: 60/123972
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: 60/127372
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/131271
; PRIOR FILING DATE: 1999-04-27
; PRIOR APPLICATION NUMBER: 60/133459
; PRIOR FILING DATE: 1999-05-11
; PRIOR APPLICATION NUMBER: 60/135725
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 60/135729
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 60/135750
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 60/138385
; PRIOR FILING DATE: 1999-06-09
; PRIOR APPLICATION NUMBER: 60/140653
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: 60/141037
; PRIOR FILING DATE: 1999-06-23
; PRIOR APPLICATION NUMBER: 60/144732
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 60/144758
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 60/144790
```



```

PRIOR FILING DATE: 1999-07-20
PRIOR APPLICATION NUMBER: 60/145228
PRIOR FILING DATE: 1999-07-20
PRIOR APPLICATION NUMBER: 60/145698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: 60/146222
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: 60/146843
PRIOR FILING DATE: 1999-08-03
PRIOR APPLICATION NUMBER: 60/148188
PRIOR FILING DATE: 1999-08-10
PRIOR APPLICATION NUMBER: 60/148513
PRIOR FILING DATE: 1999-08-12
PRIOR APPLICATION NUMBER: 60/149327
PRIOR FILING DATE: 1999-08-17
PRIOR APPLICATION NUMBER: 60/149395
PRIOR FILING DATE: 1999-08-17
PRIOR APPLICATION NUMBER: 60/150114
PRIOR FILING DATE: 1999-08-20
PRIOR APPLICATION NUMBER: 60/151700
PRIOR FILING DATE: 1999-08-31
PRIOR APPLICATION NUMBER: 60/151734
PRIOR FILING DATE: 1999-08-31
PRIOR APPLICATION NUMBER: 60/162506
PRIOR FILING DATE: 1999-10-29
PRIOR APPLICATION NUMBER: 60/170262
PRIOR FILING DATE: 1999-12-09
PRIOR APPLICATION NUMBER: 60/177118
PRIOR FILING DATE: 2000-01-20
PRIOR APPLICATION NUMBER: 60/179851
PRIOR FILING DATE: 2000-02-02
PRIOR APPLICATION NUMBER: 60/180921
PRIOR FILING DATE: 2000-02-08
PRIOR APPLICATION NUMBER: 60/187202
PRIOR FILING DATE: 2000-03-03
PRIOR APPLICATION NUMBER: 60/198587
PRIOR FILING DATE: 2000-04-18
PRIOR APPLICATION NUMBER: 60/199614
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 60/206330
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/206368
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/209832
PRIOR FILING DATE: 2000-06-05
PRIOR APPLICATION NUMBER: 60/218371
PRIOR FILING DATE: 2000-07-13
PRIOR APPLICATION NUMBER: 60/222695
PRIOR FILING DATE: 2000-08-02
PRIOR APPLICATION NUMBER: 60/229896
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/230621
PRIOR FILING DATE: 2000-09-05
PRIOR APPLICATION NUMBER: 60/232887
PRIOR FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: 60/235147
PRIOR FILING DATE: 2000-09-22
PRIOR APPLICATION NUMBER: 60/261878
PRIOR FILING DATE: 2001-01-12
PRIOR APPLICATION NUMBER: 60/261910
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: 60/261939
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: 60/262150
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: 60/264395
PRIOR FILING DATE: 2001-01-25
PRIOR APPLICATION NUMBER: 60/266421
PRIOR FILING DATE: 2001-02-02
PRIOR APPLICATION NUMBER: 60/267623
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/274399
PRIOR FILING DATE: 2001-03-09

```

```

PRIOR APPLICATION NUMBER: 60/280982
PRIOR FILING DATE: 2001-04-03
PRIOR APPLICATION NUMBER: 60/282129
PRIOR FILING DATE: 2001-04-04
PRIOR APPLICATION NUMBER: 60/282199
PRIOR FILING DATE: 2001-04-04
PRIOR APPLICATION NUMBER: 60/290589
PRIOR FILING DATE: 2001-05-09
PRIOR APPLICATION NUMBER: 09/180997
PRIOR FILING DATE: 1998-11-19
PRIOR APPLICATION NUMBER: 09/267213
PRIOR FILING DATE: 1999-03-12
PRIOR APPLICATION NUMBER: 09/380137
PRIOR FILING DATE: 1999-08-25
PRIOR APPLICATION NUMBER: 09/380138
PRIOR FILING DATE: 1999-08-25
PRIOR APPLICATION NUMBER: 09/403297
PRIOR FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: 09/423741
PRIOR FILING DATE: 1999-11-10
PRIOR APPLICATION NUMBER: 09/709238
PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 09/802706
PRIOR FILING DATE: 2001-03-09
PRIOR APPLICATION NUMBER: 09/872035
PRIOR FILING DATE: 2001-06-01
PRIOR APPLICATION NUMBER: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 09/924419
PRIOR FILING DATE: 2001-08-06
PRIOR APPLICATION NUMBER: 09/927796
PRIOR FILING DATE: 2001-08-09
PRIOR APPLICATION NUMBER: 09/929404
PRIOR FILING DATE: 2001-08-13
PRIOR APPLICATION NUMBER: 09/931836
PRIOR FILING DATE: 2001-08-16
PRIOR APPLICATION NUMBER: 09/941992
PRIOR FILING DATE: 2001-08-28
PRIOR APPLICATION NUMBER: 09/946374
PRIOR FILING DATE: 2001-09-04
PRIOR APPLICATION NUMBER: 10/001054
PRIOR FILING DATE: 2001-11-30
PRIOR APPLICATION NUMBER: 10/052586
PRIOR FILING DATE: 2002-01-15
PRIOR APPLICATION NUMBER: 10/081056
PRIOR FILING DATE: 2002-02-20
PRIOR APPLICATION NUMBER: 10/119480
PRIOR FILING DATE: 2002-04-09

Alignment Scores:
Pred. No.: 0.0459
Score: 12.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 1.09%
DB: 9
Gaps: 0

US-09-397-967-16 (1-1099) x US-10-238-183-103 (1-3583)
Oy 941 ValHSARGsPLeuAlaIAIAIAgAnIleLeuVal 952
DB 2633 GTTCATCGAGACCTGAGCGGCTCGAATATATCTGTC 2668

RESULT 26
US-10-238-283-103
Sequence 103, Application US/10238283
Publication No. US20030073190A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin
APPLICANT: Eaton, Dan
APPLICANT: Filvaroff, Ellen
APPLICANT: Goddard, Audrey
APPLICANT: Grimaldi, J. Christopher

```

```

: APPLICANT: Gurney,Austin
: APPLICANT: Smith,Victoria
: APPLICANT: Stephan,Jean-Phillippe
: APPLICANT: Watanabe,Colin
: APPLICANT: Wood,William
: APPLICANT: Zhang,Zemin
: APPLICANT: Fong,Sherman
: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
: FILE REFERENCE: P3630R1C15
: CURRENT APPLICATION NUMBER: US/10/238,283
: PRIOR FILING DATE: 2002-09-09
: PRIOR APPLICATION NUMBER: 10/197942
: PRIOR FILING DATE: 2002-07-18
: PRIOR APPLICATION NUMBER: 60/059114
: PRIOR FILING DATE: 1997-09-17
: PRIOR APPLICATION NUMBER: 60/063046
: PRIOR FILING DATE: 1997-10-24
: PRIOR APPLICATION NUMBER: 60/065027
: PRIOR FILING DATE: 1997-11-10
: PRIOR APPLICATION NUMBER: 60/079689
: PRIOR FILING DATE: 1998-03-27
: PRIOR APPLICATION NUMBER: 60/086478
: PRIOR FILING DATE: 1998-05-22
: PRIOR APPLICATION NUMBER: 60/087607
: PRIOR FILING DATE: 1998-06-02
: PRIOR APPLICATION NUMBER: 60/089801
: PRIOR FILING DATE: 1998-06-18
: PRIOR APPLICATION NUMBER: 60/090557
: PRIOR FILING DATE: 1998-06-24
: PRIOR APPLICATION NUMBER: 60/090689
: PRIOR FILING DATE: 1998-06-25
: Remaining Prior Application data removed - See File Wrapper or PALM.
: NUMBER OF SEQ ID NOS: 116
: SEQ ID NO 103
: LENGTH: 3583
: TYPE: DNA
: ORGANISM: Homo Sapien
: US-10-238-283-103

Alignment Scores:
Pred. No.: 0.0459 Length: 3583
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
Gaps: 0
DB: 9

US-09-397-967-16 (1-1099) x US-10-238-283-103 (1-3583)
QY 941 VALHISARGASPLEUALAALARGSNILEUVAL 952
DB 2633 GTTCATCGAGACTTAGCGGCTCGAATATACTGCTC 2668

RESULT 27
US-10-238-370-103
: Sequence 103, Application US/10238370
: Publication No. US20030073191A1
: GENERAL INFORMATION:
: APPLICANT: Baker,Kevin
: APPLICANT: Eaton,Dan
: APPLICANT: Filvaroff,Ellen
: APPLICANT: Goddard,Audrey
: APPLICANT: Grimaldi,J. Christopher
: APPLICANT: Gurney,Austin
: APPLICANT: Smith,Victoria
: APPLICANT: Stephan,Jean-Phillippe
: APPLICANT: Watanabe,Colin
: APPLICANT: Wood,William
: APPLICANT: Zhang,Zemin
: APPLICANT: Fong,Sherman
: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
: FILE REFERENCE: P3630R1C15
: CURRENT APPLICATION NUMBER: US/10/238,283
: PRIOR FILING DATE: 2002-09-09
: PRIOR APPLICATION NUMBER: 10/197942
: PRIOR FILING DATE: 2002-07-18
: PRIOR APPLICATION NUMBER: 60/059114
: PRIOR FILING DATE: 1997-09-17
: PRIOR APPLICATION NUMBER: 60/063046
: PRIOR FILING DATE: 1997-10-24
: PRIOR APPLICATION NUMBER: 60/065027
: PRIOR FILING DATE: 1997-11-10
: PRIOR APPLICATION NUMBER: 60/079689
: PRIOR FILING DATE: 1998-03-27
: PRIOR APPLICATION NUMBER: 60/086478
: PRIOR FILING DATE: 1998-05-22
: PRIOR APPLICATION NUMBER: 60/087607
: PRIOR FILING DATE: 1998-06-02
: PRIOR APPLICATION NUMBER: 60/089801
: PRIOR FILING DATE: 1998-06-18
: PRIOR APPLICATION NUMBER: 60/090557
: PRIOR FILING DATE: 1998-06-24
: PRIOR APPLICATION NUMBER: 60/090689
: PRIOR FILING DATE: 1998-06-25
: Remaining Prior Application data removed - See File Wrapper or PALM.
: NUMBER OF SEQ ID NOS: 116
: SEQ ID NO 103
: LENGTH: 3583
: TYPE: DNA
: ORGANISM: Homo Sapien
: US-10-238-370-103

Alignment Scores:
Pred. No.: 0.0459 Length: 3583
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
Gaps: 0
DB: 9

US-09-397-967-16 (1-1099) x US-10-238-370-103 (1-3583)
QY 941 VALHISARGASPLEUALAALARGSNILEUVAL 952
DB 2633 GTTCATCGAGACTTAGCGGCTCGAATATACTGCTC 2668

RESULT 28
US-10-245-055-103
: Sequence 103, Application US/10245055
: Publication No. US20030073192A1
: GENERAL INFORMATION:
: APPLICANT: Baker,Kevin
: APPLICANT: Eaton,Dan
: APPLICANT: Filvaroff,Ellen
: APPLICANT: Goddard,Audrey
: APPLICANT: Grimaldi,J. Christopher
: APPLICANT: Gurney,Austin
: APPLICANT: Smith,Victoria
: APPLICANT: Stephan,Jean-Phillippe
: APPLICANT: Watanabe,Colin
: APPLICANT: Wood,William
: APPLICANT: Zhang,Zemin
: APPLICANT: Fong,Sherman
: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
: FILE REFERENCE: P3630R1C88
: CURRENT APPLICATION NUMBER: US/10/245,055
: PRIOR FILING DATE: 2002-09-16
: PRIOR APPLICATION NUMBER: 10/197942
: PRIOR FILING DATE: 2002-07-18
: PRIOR APPLICATION NUMBER: 60/059114
: PRIOR FILING DATE: 1997-09-17
: PRIOR APPLICATION NUMBER: 60/063046
: PRIOR FILING DATE: 1997-10-24
```

```

: FILE REFERENCE: P3630R1C10
: CURRENT APPLICATION NUMBER: US/10/238,370
: PRIOR FILING DATE: 2002-09-09
: PRIOR APPLICATION NUMBER: 10/197942
: PRIOR FILING DATE: 2002-07-18
: PRIOR APPLICATION NUMBER: 60/059114
: PRIOR FILING DATE: 1997-09-17
: PRIOR APPLICATION NUMBER: 60/063046
: PRIOR FILING DATE: 1997-10-24
: PRIOR APPLICATION NUMBER: 60/065027
: PRIOR FILING DATE: 1997-11-10
: PRIOR APPLICATION NUMBER: 60/079689
: PRIOR FILING DATE: 1998-03-27
: PRIOR APPLICATION NUMBER: 60/086478
: PRIOR FILING DATE: 1998-05-22
: PRIOR APPLICATION NUMBER: 60/087607
: PRIOR FILING DATE: 1998-06-02
: PRIOR APPLICATION NUMBER: 60/089801
: PRIOR FILING DATE: 1998-06-18
: PRIOR APPLICATION NUMBER: 60/090557
: PRIOR FILING DATE: 1998-06-24
: PRIOR APPLICATION NUMBER: 60/090689
: PRIOR FILING DATE: 1998-06-25
: Remaining Prior Application data removed - See File Wrapper or PALM.
: NUMBER OF SEQ ID NOS: 116
: SEQ ID NO 103
: LENGTH: 3583
: TYPE: DNA
: ORGANISM: Homo Sapien
: US-10-238-370-103

Alignment Scores:
Pred. No.: 0.0459 Length: 3583
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
Gaps: 0
DB: 9

US-09-397-967-16 (1-1099) x US-10-238-370-103 (1-3583)
QY 941 VALHISARGASPLEUALAALARGSNILEUVAL 952
DB 2633 GTTCATCGAGACTTAGCGGCTCGAATATACTGCTC 2668

RESULT 28
US-10-245-055-103
: Sequence 103, Application US/10245055
: Publication No. US20030073192A1
: GENERAL INFORMATION:
: APPLICANT: Baker,Kevin
: APPLICANT: Eaton,Dan
: APPLICANT: Filvaroff,Ellen
: APPLICANT: Goddard,Audrey
: APPLICANT: Grimaldi,J. Christopher
: APPLICANT: Gurney,Austin
: APPLICANT: Smith,Victoria
: APPLICANT: Stephan,Jean-Phillippe
: APPLICANT: Watanabe,Colin
: APPLICANT: Wood,William
: APPLICANT: Zhang,Zemin
: APPLICANT: Fong,Sherman
: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
: FILE REFERENCE: P3630R1C88
: CURRENT APPLICATION NUMBER: US/10/245,055
: PRIOR FILING DATE: 2002-09-16
: PRIOR APPLICATION NUMBER: 10/197942
: PRIOR FILING DATE: 2002-07-18
: PRIOR APPLICATION NUMBER: 60/059114
: PRIOR FILING DATE: 1997-09-17
: PRIOR APPLICATION NUMBER: 60/063046
: PRIOR FILING DATE: 1997-10-24
```

```
;; PRIOR APPLICATION NUMBER: 60/065027
;; PRIOR FILING DATE: 1997-11-10
;; PRIOR APPLICATION NUMBER: 60/079689
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/086478
;; PRIOR FILING DATE: 1998-05-22
;; PRIOR APPLICATION NUMBER: 60/087607
;; PRIOR FILING DATE: 1998-06-02
;; PRIOR APPLICATION NUMBER: 60/089801
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/090557
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090689
;; PRIOR FILING DATE: 1998-06-25
;; Remaining Prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 116
;; SEQ ID NO 103
;; LENGTH: 3583
;; TYPE: DNA
;; ORGANISM: Homo Sapien
US-10-245-055-103

Alignment Scores:
Pred. No.: 0.0459      Length: 3583
Score: 12.00          Matches: 12
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 1.09%      Indels: 0
DB: 9      Gaps: 0

US-09-397-967-16 (1-1099) x US-10-245-055-103 (1-3583)

QY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
Db 2633 GTTCATCGAGACCTAGCGGCTCGGAATATCTGTC 2668

RESULT 29
US-10-245-147-103
;; Sequence 103, Application US/10245147
;; Publication No. US20030073193A1
;; GENERAL INFORMATION:
;; APPLICANT: Baker, Kevin
;; APPLICANT: Baton, Dan
;; APPLICANT: Filvaroff, Ellen
;; APPLICANT: Goddard, Audrey
;; APPLICANT: Grimaldi, J. Christopher
;; APPLICANT: Gurney, Austin
;; APPLICANT: Smith, Victoria
;; APPLICANT: Stephan, Jean-Philippe
;; APPLICANT: Watanabe, Colin
;; APPLICANT: Wood, William
;; APPLICANT: Zhang, Zemin
;; APPLICANT: Fong, Sherman
;; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;; FILE REFERENCE: P3630R1C72
;; CURRENT APPLICATION NUMBER: US/10/245,147
;; PRIOR FILING DATE: 2002-09-16
;; PRIOR APPLICATION NUMBER: 10/197942
;; PRIOR FILING DATE: 2002-07-18
;; PRIOR APPLICATION NUMBER: 60/059114
;; PRIOR FILING DATE: 1997-09-17
;; PRIOR APPLICATION NUMBER: 60/063046
;; PRIOR FILING DATE: 1997-10-24
;; PRIOR APPLICATION NUMBER: 60/065027
;; PRIOR FILING DATE: 1997-11-10
;; PRIOR APPLICATION NUMBER: 60/079689
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/086478
;; PRIOR FILING DATE: 1998-05-22
;; PRIOR APPLICATION NUMBER: 60/087607
;; PRIOR FILING DATE: 1998-06-02
;; PRIOR APPLICATION NUMBER: 60/089801
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/090557
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090689
;; Remaining Prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 116
;; SEQ ID NO 103
;; LENGTH: 3583
```

```
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/090557
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090689
;; PRIOR FILING DATE: 1998-06-25
;; Remaining Prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 116
;; SEQ ID NO 103
;; LENGTH: 3583
US-10-245-147-103

Alignment Scores:
Pred. No.: 0.0459      Length: 3583
Score: 12.00          Matches: 12
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 1.09%      Indels: 0
DB: 9      Gaps: 0

US-09-397-967-16 (1-1099) x US-10-245-147-103 (1-3583)

QY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
Db 2633 GTTCATCGAGACCTAGCGGCTCGGAATATCTGTC 2668

RESULT 30
US-10-245-730-103
;; Sequence 103, Application US/10245730
;; Publication No. US20030073194A1
;; GENERAL INFORMATION:
;; APPLICANT: Baker, Kevin
;; APPLICANT: Baton, Dan
;; APPLICANT: Filvaroff, Ellen
;; APPLICANT: Goddard, Audrey
;; APPLICANT: Grimaldi, J. Christopher
;; APPLICANT: Gurney, Austin
;; APPLICANT: Smith, Victoria
;; APPLICANT: Stephan, Jean-Philippe
;; APPLICANT: Watanabe, Colin
;; APPLICANT: Wood, William
;; APPLICANT: Zhang, Zemin
;; APPLICANT: Fong, Sherman
;; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;; FILE REFERENCE: P3630R1C85
;; CURRENT APPLICATION NUMBER: US/10/245,730
;; PRIOR FILING DATE: 2002-09-16
;; PRIOR APPLICATION NUMBER: 10/197942
;; PRIOR FILING DATE: 2002-07-18
;; PRIOR APPLICATION NUMBER: 60/059114
;; PRIOR FILING DATE: 1997-09-17
;; PRIOR APPLICATION NUMBER: 60/063046
;; PRIOR FILING DATE: 1997-10-24
;; PRIOR APPLICATION NUMBER: 60/065027
;; PRIOR FILING DATE: 1997-11-10
;; PRIOR APPLICATION NUMBER: 60/079689
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/086478
;; PRIOR FILING DATE: 1998-05-22
;; PRIOR APPLICATION NUMBER: 60/087607
;; PRIOR FILING DATE: 1998-06-02
;; PRIOR APPLICATION NUMBER: 60/089801
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/090557
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090689
;; Remaining Prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 116
;; SEQ ID NO 103
;; LENGTH: 3583
```

```

; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-245-730-103

Alignment Scores:
Pred. No.: 0.0459 Length: 3583
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 9 Gaps: 0

US-09-397-967-16 (1-1099) x US-10-245-730-103 (1-3583)

QY 941 VALHISARGASPLEUALAALARGASNIILEUVAL 952
DB 2633 GTTCATCGAGACCTAGCGGCTCGAATATACGTGTC 2668

RESULT 31
US-10-245-739-103
; Sequence 103, Application US/10245739
; Publication No. US20030073195A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Eaton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe
; APPLICANT: Watanabe, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3630R1C86
; CURRENT APPLICATION NUMBER: US/10/245,739
; CURRENT FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 103
; LENGTH: 3583
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-245-739-103

Alignment Scores:
Pred. No.: 0.0459 Length: 3583
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 9 Gaps: 0

```

```

Query Match: 1.09% Indels: 0
DB: 9 Gaps: 0

US-09-397-967-16 (1-1099) x US-10-245-739-103 (1-3583)

QY 941 VALHISARGASPLEUALAALARGASNIILEUVAL 952
DB 2633 GTTCATCGAGACCTAGCGGCTCGAATATACGTGTC 2668

RESULT 32
US-10-246-210-103
; Sequence 103, Application US/10246210
; Publication No. US20030073196A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Eaton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe
; APPLICANT: Watanabe, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3630R1C121
; CURRENT APPLICATION NUMBER: US/10/246,210
; CURRENT FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 103
; LENGTH: 3583
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-246-210-103

Alignment Scores:
Pred. No.: 0.0459 Length: 3583
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 9 Gaps: 0

US-09-397-967-16 (1-1099) x US-10-246-210-103 (1-3583)

QY 941 VALHISARGASPLEUALAALARGASNIILEUVAL 952
DB 2633 GTTCATCGAGACCTAGCGGCTCGAATATACGTGTC 2668

```


;; TITLE OF INVENTION: Treating Colon Cancer
;; FILE REFERENCE: DEX-0211
;; CURRENT APPLICATION NUMBER: US/09/895,652
;; CURRENT FILING DATE: 2001-06-28
;; PRIOR APPLICATION NUMBER: 60/214,515
;; PRIOR FILING DATE: 2000-06-28
;; NUMBER OF SEQ ID NOS: 39
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 1
;; LENGTH: 4698
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: unsure
;; LOCATION: (3921)
US-09-895-652-1

Alignment Scores:
Pred. No.: 0.0585 Length: 4698
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: Gaps: 0

US-09-397-967-16 (1-1099) x US-09-895-652-1 (1-4698)

OY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
|||||
Db 2237 GTTCACCGTGGCTGGCTGCCGCAACATCTCTGTC 2272

RESULT 37
US-09-205-658-11

;; Sequence 11, Application US/09205658
;; Patent No. US20010029617A1
;; GENERAL INFORMATION:
;; APPLICANT: Ruvkun, Gary
;; APPLICANT: Ogg, Scott
;; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC TOOLS FOR
;; TITLE OF INVENTION: IMPAIRED GLUCOSE TOLERANCE CONDITIONS
;; FILE REFERENCE: 00786/351004
;; CURRENT APPLICATION NUMBER: US/09/205,658
;; CURRENT FILING DATE: 1998-12-03
;; EARLIER APPLICATION NUMBER: 08/857,076
;; EARLIER FILING DATE: 1997-05-15
;; EARLIER APPLICATION NUMBER: 08/888,534
;; EARLIER FILING DATE: 1997-07-07
;; EARLIER APPLICATION NUMBER: US98/10080
;; EARLIER FILING DATE: 1998-05-15
;; NUMBER OF SEQ ID NOS: 328
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 11
;; LENGTH: 5816
;; TYPE: DNA
;; ORGANISM: Caenorhabditis elegans
;; FEATURE:
;; NAME/KEY: misc.feature
;; LOCATION: (1)...(5816)
;; OTHER INFORMATION: n = A,T,C or G
US-09-205-658-11

Alignment Scores:
Pred. No.: 0.0708 Length: 5816
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: Gaps: 0

US-09-397-967-16 (1-1099) x US-09-205-658-11 (1-5816)

OY 1003 SerAspValTTPSerPheGlyValValLeuTyrglu 1014
|||||
US-09-397-967-16 (1-1099) x US-09-205-658-11 (1-5816)

Db 4455 TCTGATGTTTGAGCTTCGAGTGTCTCTATGAA 4490

RESULT 38
US-09-844-353A-11

;; Sequence 11, Application US/09844353A
;; Patent No. US20020037585A1
;; GENERAL INFORMATION:
;; APPLICANT: Ruvkun, Gary
;; APPLICANT: Kimura, Gary
;; APPLICANT: Patterson, Garth
;; APPLICANT: Ogg, Scott
;; APPLICANT: Paradis, Suzanne
;; APPLICANT: Tissenbaum, Heidi
;; APPLICANT: Morris, Jason
;; APPLICANT: Kowek, Allison
;; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC TOOLS FOR
;; TITLE OF INVENTION: IMPAIRED GLUCOSE TOLERANCE CONDITIONS
;; FILE REFERENCE: 00786/351005
;; CURRENT APPLICATION NUMBER: US/09/844,353A
;; CURRENT FILING DATE: 2001-04-27
;; PRIOR APPLICATION NUMBER: US 08/857,076
;; PRIOR FILING DATE: 1997-05-15
;; NUMBER OF SEQ ID NOS: 114
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 11
;; LENGTH: 5816
;; TYPE: DNA
;; ORGANISM: Caenorhabditis elegans
;; FEATURE:
;; NAME/KEY: misc.feature
;; LOCATION: (1)...(5816)
;; OTHER INFORMATION: n = A,T,C or G
US-09-844-353A-11

Alignment Scores:
Pred. No.: 0.0708 Length: 5816
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: Gaps: 0

US-09-397-967-16 (1-1099) x US-09-844-353A-11 (1-5816)

OY 1003 SerAspValTTPSerPheGlyValValLeuTyrglu 1014
|||||
Db 4455 TCTGATGTTTGAGCTTCGAGTGTCTCTATGAA 4490

RESULT 39
US-09-982-610-7

;; Sequence 7, Application US/09982610
;; Patent No. US20020146420A1
;; GENERAL INFORMATION:
;; APPLICANT: Genentech, Inc.
;; APPLICANT: Bennett, Brian D.
;; APPLICANT: Goeddel, David
;; APPLICANT: Lee, James M.
;; APPLICANT: Matthews, William
;; APPLICANT: Tsai, Siao ping
;; APPLICANT: Wood, William I.
;; TITLE OF INVENTION: PROTEIN TYROSINE KINASE AGONIST ANTIBODIES
;; NUMBER OF SEQUENCES: 45
;; CORRESPONDENCE ADDRESS:
;; ADDRESS: Genentech, Inc.
;; STREET: 460 Point San Bruno Blvd
;; CITY: South San Francisco
;; STATE: California
;; COUNTRY: USA
;; ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS


```

GENERAL INFORMATION:
APPLICANT: Gaiger, Alexander
APPLICANT: Algate, Paul A.
APPLICANT: Mainion, Jane
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
FILE REFERENCE: 2077, 001200
CURRENT APPLICATION NUMBER: US/09/796,692
CURRENT FILING DATE: 2001-03-01
PRIORITY APPLICATION NUMBER: 60/186,126
PRIORITY FILING DATE: 2000-03-01
PRIORITY APPLICATION NUMBER: 60/190,479
PRIORITY FILING DATE: 2000-03-17
PRIORITY APPLICATION NUMBER: 60/200,545
PRIORITY FILING DATE: 2000-04-27
PRIORITY APPLICATION NUMBER: 60/200,303
PRIORITY FILING DATE: 2000-04-28
PRIORITY APPLICATION NUMBER: 60/200,779
PRIORITY FILING DATE: 2000-04-28
PRIORITY APPLICATION NUMBER: 60/200,999
PRIORITY FILING DATE: 2000-05-01
PRIORITY APPLICATION NUMBER: 60/202,084
PRIORITY FILING DATE: 2000-05-04
PRIORITY APPLICATION NUMBER: 60/206,201
PRIORITY FILING DATE: 2000-05-22
PRIORITY APPLICATION NUMBER: 60/218,950
PRIORITY FILING DATE: 2000-07-14
PRIORITY APPLICATION NUMBER: 60/222,903
PRIORITY FILING DATE: 2000-08-03
PRIORITY APPLICATION NUMBER: 60/223,416
PRIORITY FILING DATE: 2000-08-04
PRIORITY APPLICATION NUMBER: 60/223,378
PRIORITY FILING DATE: 2000-08-07
NUMBER OF SEQ ID NOS: 9597
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 8268
LENGTH: 510
TYPE: DNA
ORGANISM: Homo sapiens
US-09-796-692-8268

Alignment Scores:
Pred. No.: 0.887 Length: 510
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.91% Indels: 0
DB: 9 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-796-692-8268 (1-510)
Oy 940 CysvalHisargaspLeuAlaIaargasn 949
|||||
Db 375 TGTGTTACAGAGACTGCGCCGACGAGAAC 404

RESULT 43
US-10-101-464A-355
; Sequence 355, Application US/10101464A
; Publication No. US20030046728A1
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; APPLICANT: Higgins, Colleen M.
TITLE OF INVENTION: Compositions Isolated from Plant Cells
TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling
FILE REFERENCE: 11000,1020c2
CURRENT APPLICATION NUMBER: US/10/101,464A
CURRENT FILING DATE: 2002-03-18
PRIORITY APPLICATION NUMBER: 09/704,302
PRIORITY FILING DATE: 2000-11-01
PRIORITY APPLICATION NUMBER: 09/228,986
PRIORITY FILING DATE: 1999-01-12
PRIORITY APPLICATION NUMBER: 60/162,866

```

```

: PRIOR FILING DATE: 1999-11-01
: PRIOR APPLICATION NUMBER: PCT/US00/00724
: PRIOR FILING DATE: 2000-01-11
: NUMBER OF SEQ ID NOS: 989
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 355
: LENGTH: 536
: TYPE: DNA
: ORGANISM: Pinus radiata
US-10-101-464A-355

Alignment Scores:
Pred. No.: 0.927 Length: 536
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.91% Indels: 0
DB: 9 Gaps: 0

US-09-397-967-16 (1-1099) x US-10-101-464A-355 (1-536)
Oy 959 LysileAlaasphegLyLeuAlaLysLeu 968
|||||
Db 329 AAATAGCGGATTTGCAITGCAAAACTT 358

RESULT 44
US-09-796-692-8255/c
: Sequence 8255, Application US/09796692
: Publication No. US20020198362A1
: GENERAL INFORMATION:
: APPLICANT: Gaiger, Alexander
: APPLICANT: Algate, Paul A.
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THER
: FILE REFERENCE: 2077. 001200
: CURRENT APPLICATION NUMBER: US/09/796,692
: PRIOR FILING DATE: 2001-03-01
: PRIOR APPLICATION NUMBER: 60/186,126
: PRIOR FILING DATE: 2000-03-01
: PRIOR APPLICATION NUMBER: 60/190,479
: PRIOR FILING DATE: 2000-03-17
: PRIOR APPLICATION NUMBER: 60/200,545
: PRIOR FILING DATE: 2000-04-27
: PRIOR APPLICATION NUMBER: 60/200,303
: PRIOR FILING DATE: 2000-04-28
: PRIOR APPLICATION NUMBER: 60/200,779
: PRIOR FILING DATE: 2000-04-28
: PRIOR APPLICATION NUMBER: 60/200,999
: PRIOR FILING DATE: 2000-05-01
: PRIOR APPLICATION NUMBER: 60/202,084
: PRIOR FILING DATE: 2000-05-04
: PRIOR APPLICATION NUMBER: 60/206,201
: PRIOR FILING DATE: 2000-05-22
: PRIOR APPLICATION NUMBER: 60/218,950
: PRIOR FILING DATE: 2000-07-14
: PRIOR APPLICATION NUMBER: 60/222,903
: PRIOR FILING DATE: 2000-08-03
: PRIOR APPLICATION NUMBER: 60/223,416
: PRIOR FILING DATE: 2000-08-04
: PRIOR APPLICATION NUMBER: 60/223,378
: PRIOR FILING DATE: 2000-08-07
: NUMBER OF SEQ ID NOS: 9597
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 8255
: LENGTH: 539
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-796-692-8255

Alignment Scores:
Pred. No.: 0.932 Length: 539
Score: 10.00 Matches: 10

```



```
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match: 0.91%               Indels: 0
DB: 9                             Gaps: 0

US-09-397-967-16 (1-1099) x US-09-796-692-8255 (1-539)
QY 940 CysValHisArgAspLeuAlaIaArgAsn 949
   |||||||
Db 420 TGTGTCACAGAGACCTGGCCGCCAGAGAC 391

RESULT 45
US-09-796-692-7883
; Sequence 7883, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; PRIOR FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7883
; LENGTH: 556
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-796-692-7883

Alignment Scores:
Pred. No.: 0.958      Length: 556
Score: 10.00          Matches: 10
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match: 0.91%      Indels: 0
DB: 9                Gaps: 0

US-09-397-967-16 (1-1099) x US-09-796-692-7883 (1-556)
QY 940 CysValHisArgAspLeuAlaIaArgAsn 949
   |||||||
Db 418 TGTGTCACAGAGACCTGGCCGCCAGAGAC 447

RESULT 46
US-09-796-692-9025
; Sequence 9025, Application US/09796692
```

```
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THER
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; PRIOR FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9025
; LENGTH: 564
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-796-692-9025

Alignment Scores:
Pred. No.: 0.97      Length: 564
Score: 10.00          Matches: 10
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match: 0.91%      Indels: 0
DB: 9                Gaps: 0

US-09-397-967-16 (1-1099) x US-09-796-692-9025 (1-564)
QY 940 CysValHisArgAspLeuAlaIaArgAsn 949
   |||||||
Db 418 TGTGTCACAGAGACCTGGCCGCCAGAGAC 447

RESULT 47
US-09-796-692-8630
; Sequence 8630, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THER
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; PRIOR FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8630
; LENGTH: 564
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-796-692-8630
```

```
;; PRIOR APPLICATION NUMBER: 60/200,545
;; PRIOR FILING DATE: 2000-04-27
;; PRIOR APPLICATION NUMBER: 60/200,303
;; PRIOR FILING DATE: 2000-04-28
;; PRIOR APPLICATION NUMBER: 60/200,779
;; PRIOR FILING DATE: 2000-04-28
;; PRIOR APPLICATION NUMBER: 60/200,999
;; PRIOR FILING DATE: 2000-05-01
;; PRIOR APPLICATION NUMBER: 60/202,084
;; PRIOR FILING DATE: 2000-05-04
;; PRIOR APPLICATION NUMBER: 60/206,201
;; PRIOR FILING DATE: 2000-05-22
;; PRIOR APPLICATION NUMBER: 60/218,950
;; PRIOR FILING DATE: 2000-07-14
;; PRIOR APPLICATION NUMBER: 60/222,903
;; PRIOR FILING DATE: 2000-08-03
;; PRIOR APPLICATION NUMBER: 60/223,416
;; PRIOR FILING DATE: 2000-08-04
;; PRIOR APPLICATION NUMBER: 60/223,378
;; PRIOR FILING DATE: 2000-08-07
;; NUMBER OF SEQ ID NOS: 9597
;; SOFTWARE: FASTSEQ for Windows Version 3.0
;; SEQ ID NO 8630
;; LENGTH: 589
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: unsure
;; LOCATION: (50)
;; OTHER INFORMATION: n-A,T,C or G
;; NAME/KEY: (231)
;; LOCATION: (231)
;; OTHER INFORMATION: n-A,T,C or G
US-09-796-692-8630

Alignment Scores:
Pred. No.: 1.01 Length: 589
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.91% Indels: 0
DB: 9 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-796-692-8630 (1-589)
Oy 940 CysValHisArgAspLeuAlaIlaArgAsn 949
Db 418 TGTGTTCAAGAGACCTGGCCGCGCAGAAC 447

RESULT 48
US-09-796-692-7976
;; Sequence 7976, Application US/09796692
;; Publication No. US20020198362A1
;; GENERAL INFORMATION:
;; APPLICANT: Gaiger, Alexander
;; APPLICANT: Algate, Paul A.
;; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
;; TITLE OF INVENTION: HEMATOLOGICAL MALIGNANCIES
;; FILE REFERENCE: 2077.001200
;; PRIOR APPLICATION NUMBER: US/09/796,692
;; CURRENT FILING DATE: 2001-03-01
;; PRIOR APPLICATION NUMBER: 60/186,126
;; PRIOR FILING DATE: 2000-03-01
;; PRIOR APPLICATION NUMBER: 60/190,479
;; PRIOR FILING DATE: 2000-03-17
;; PRIOR APPLICATION NUMBER: 60/200,545
;; PRIOR FILING DATE: 2000-04-27
;; PRIOR APPLICATION NUMBER: 60/200,303
;; PRIOR FILING DATE: 2000-04-28
;; PRIOR APPLICATION NUMBER: 60/200,779
;; PRIOR FILING DATE: 2000-04-28
;; PRIOR APPLICATION NUMBER: 60/200,999
;; PRIOR FILING DATE: 2000-04-28
;; PRIOR APPLICATION NUMBER: 60/200,999
```

```
;; PRIOR FILING DATE: 2000-05-01
;; PRIOR APPLICATION NUMBER: 60/202,084
;; PRIOR FILING DATE: 2000-05-04
;; PRIOR APPLICATION NUMBER: 60/206,201
;; PRIOR FILING DATE: 2000-05-22
;; PRIOR APPLICATION NUMBER: 60/218,950
;; PRIOR FILING DATE: 2000-07-14
;; PRIOR APPLICATION NUMBER: 60/222,903
;; PRIOR FILING DATE: 2000-08-03
;; PRIOR APPLICATION NUMBER: 60/223,416
;; PRIOR FILING DATE: 2000-08-04
;; PRIOR APPLICATION NUMBER: 60/223,378
;; PRIOR FILING DATE: 2000-08-07
;; NUMBER OF SEQ ID NOS: 9597
;; SOFTWARE: FASTSEQ for Windows Version 3.0
;; SEQ ID NO 7976
;; LENGTH: 597
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-09-796-692-7976

Alignment Scores:
Pred. No.: 1.02 Length: 597
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.91% Indels: 0
DB: 9 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-796-692-7976 (1-597)
Oy 940 CysValHisArgAspLeuAlaIlaArgAsn 949
Db 418 TGTGTTCAAGAGACCTGGCCGCGCAGAAC 447

RESULT 49
US-09-796-692-9010
;; Sequence 9010, Application US/09796692
;; Publication No. US20020198362A1
;; GENERAL INFORMATION:
;; APPLICANT: Gaiger, Alexander
;; APPLICANT: Algate, Paul A.
;; APPLICANT: Mannion, Jane
;; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THER
;; TITLE OF INVENTION: HEMATOLOGICAL MALIGNANCIES
;; FILE REFERENCE: 2077.001200
;; CURRENT FILING DATE: US/09/796,692
;; PRIOR APPLICATION NUMBER: 60/186,126
;; PRIOR FILING DATE: 2000-03-01
;; PRIOR APPLICATION NUMBER: 60/190,479
;; PRIOR FILING DATE: 2000-03-17
;; PRIOR APPLICATION NUMBER: 60/200,545
;; PRIOR FILING DATE: 2000-04-27
;; PRIOR APPLICATION NUMBER: 60/200,303
;; PRIOR FILING DATE: 2000-04-28
;; PRIOR APPLICATION NUMBER: 60/200,779
;; PRIOR FILING DATE: 2000-04-28
;; PRIOR APPLICATION NUMBER: 60/200,999
;; PRIOR FILING DATE: 2000-05-01
;; PRIOR APPLICATION NUMBER: 60/202,084
;; PRIOR FILING DATE: 2000-05-04
;; PRIOR APPLICATION NUMBER: 60/206,201
;; PRIOR FILING DATE: 2000-05-22
;; PRIOR APPLICATION NUMBER: 60/218,950
;; PRIOR FILING DATE: 2000-07-14
;; PRIOR APPLICATION NUMBER: 60/222,903
;; PRIOR FILING DATE: 2000-08-03
;; PRIOR APPLICATION NUMBER: 60/223,416
;; PRIOR FILING DATE: 2000-08-04
;; PRIOR APPLICATION NUMBER: 60/223,378
;; PRIOR FILING DATE: 2000-08-07
;; NUMBER OF SEQ ID NOS: 9597
```

```

: SOFTWARE: FASTSEQ for Windows Version 3.0
: SEQ ID NO 9010
: LENGTH: 597
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-796-692-9010

Alignment Scores:
Pred. No.: 1.02 Length: 597
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.91% Indels: 0
DB: 9 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-796-692-9010 (1-597)

Oy 940 CysValHisArgAspLeuAlaIArgAsn 949
Db 418 TGTGTTCAAGAGACCTGGCCGCCAGGAAC 447

RESULT 50
US-09-796-692-8273
: Sequence 8273, Application US/09796692
: Publication No. US20020198362A1
: GENERAL INFORMATION:
: APPLICANT: Gaiger, Alexander
: APPLICANT: Algate, Paul A.
: APPLICANT: Mannion, Jane
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
: FILE REFERENCE: 2077.001200
: CURRENT FILING DATE: 2001-03-01
: PRIOR FILING DATE: 2000-03-01
: PRIOR APPLICATION NUMBER: 60/186,126
: PRIOR FILING DATE: 2000-03-01
: PRIOR APPLICATION NUMBER: 60/190,479
: PRIOR FILING DATE: 2000-03-17
: PRIOR APPLICATION NUMBER: 60/200,545
: PRIOR FILING DATE: 2000-04-27
: PRIOR APPLICATION NUMBER: 60/200,303
: PRIOR FILING DATE: 2000-04-28
: PRIOR APPLICATION NUMBER: 60/200,779
: PRIOR FILING DATE: 2000-04-28
: PRIOR APPLICATION NUMBER: 60/200,999
: PRIOR FILING DATE: 2000-05-01
: PRIOR APPLICATION NUMBER: 60/202,084
: PRIOR FILING DATE: 2000-05-04
: PRIOR APPLICATION NUMBER: 60/206,201
: PRIOR FILING DATE: 2000-05-22
: PRIOR APPLICATION NUMBER: 60/218,950
: PRIOR FILING DATE: 2000-07-14
: PRIOR APPLICATION NUMBER: 60/222,903
: PRIOR FILING DATE: 2000-08-03
: PRIOR APPLICATION NUMBER: 60/223,416
: PRIOR FILING DATE: 2000-08-04
: PRIOR APPLICATION NUMBER: 60/223,378
: PRIOR FILING DATE: 2000-08-07
: NUMBER OF SEQ ID NOS: 9597
: SOFTWARE: FASTSEQ for Windows Version 3.0
: SEQ ID NO 8273
: LENGTH: 601
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: unsure
: LOCATION: (586)
: OTHER INFORMATION: n=A,T,C or G
US-09-796-692-8273

Alignment Scores:
Pred. No.: 1.03 Length: 601
Score: 10.00 Matches: 10
```

```

Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.91% Indels: 0
DB: 9 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-796-692-8273 (1-601)

Oy 940 CysValHisArgAspLeuAlaIArgAsn 949
Db 418 TGTGTTCAAGAGACCTGGCCGCCAGGAAC 447

RESULT 51
US-09-796-692-8047
: Sequence 8047, Application US/09796692
: Publication No. US20020198362A1
: GENERAL INFORMATION:
: APPLICANT: Gaiger, Alexander
: APPLICANT: Algate, Paul A.
: APPLICANT: Mannion, Jane
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THER
: FILE REFERENCE: 2077.001200
: CURRENT FILING DATE: 2001-03-01
: PRIOR FILING DATE: 2000-03-01
: PRIOR APPLICATION NUMBER: 60/186,126
: PRIOR FILING DATE: 2000-03-01
: PRIOR APPLICATION NUMBER: 60/190,479
: PRIOR FILING DATE: 2000-03-17
: PRIOR APPLICATION NUMBER: 60/200,545
: PRIOR FILING DATE: 2000-04-27
: PRIOR APPLICATION NUMBER: 60/200,303
: PRIOR FILING DATE: 2000-04-28
: PRIOR APPLICATION NUMBER: 60/200,779
: PRIOR FILING DATE: 2000-04-28
: PRIOR APPLICATION NUMBER: 60/200,999
: PRIOR FILING DATE: 2000-05-01
: PRIOR APPLICATION NUMBER: 60/202,084
: PRIOR FILING DATE: 2000-05-04
: PRIOR APPLICATION NUMBER: 60/206,201
: PRIOR FILING DATE: 2000-05-22
: PRIOR APPLICATION NUMBER: 60/218,950
: PRIOR FILING DATE: 2000-07-14
: PRIOR APPLICATION NUMBER: 60/222,903
: PRIOR FILING DATE: 2000-08-03
: PRIOR APPLICATION NUMBER: 60/223,416
: PRIOR FILING DATE: 2000-08-04
: PRIOR APPLICATION NUMBER: 60/223,378
: PRIOR FILING DATE: 2000-08-07
: NUMBER OF SEQ ID NOS: 9597
: SOFTWARE: FASTSEQ for Windows Version 3.0
: SEQ ID NO 8047
: LENGTH: 605
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: unsure
: LOCATION: (542)
: OTHER INFORMATION: n=A,T,C or G
: NAME/KEY: unsure
: LOCATION: (592)
: OTHER INFORMATION: n=A,T,C or G
US-09-796-692-8047

Alignment Scores:
Pred. No.: 1.03 Length: 605
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.91% Indels: 0
DB: 9 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-796-692-8047 (1-605)
```

```
Oy 940 CysValHisArgAspLeuAlaAlaArgasn 949
|||||
Db 418 TGTGTTCAAGAGACCTGCGCCGACGAGAAC 447

RESULT 52
US-09-796-692-8927
: Sequence 8927, Application US/09796692
: Publication No. US20020198362A1
: GENERAL INFORMATION:
: APPLICANT: Gaiger, Alexander
: APPLICANT: Aigate, Paul A.
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
: FILE REFERENCE: 2077.001200
: CURRENT APPLICATION NUMBER: US/09/796,692
: CURRENT FILING DATE: 2001-03-01
: PRIOR APPLICATION NUMBER: 60/186,126
: PRIOR FILING DATE: 2000-03-01
: PRIOR APPLICATION NUMBER: 60/190,479
: PRIOR FILING DATE: 2000-03-17
: PRIOR APPLICATION NUMBER: 60/200,545
: PRIOR FILING DATE: 2000-04-27
: PRIOR APPLICATION NUMBER: 60/200,303
: PRIOR FILING DATE: 2000-04-28
: PRIOR APPLICATION NUMBER: 60/200,779
: PRIOR FILING DATE: 2000-04-28
: PRIOR APPLICATION NUMBER: 60/200,999
: PRIOR FILING DATE: 2000-05-01
: PRIOR APPLICATION NUMBER: 60/202,084
: PRIOR FILING DATE: 2000-05-04
: PRIOR APPLICATION NUMBER: 60/206,201
: PRIOR FILING DATE: 2000-05-22
: PRIOR APPLICATION NUMBER: 60/218,950
: PRIOR FILING DATE: 2000-07-14
: PRIOR APPLICATION NUMBER: 60/222,903
: PRIOR FILING DATE: 2000-08-03
: PRIOR APPLICATION NUMBER: 60/223,416
: PRIOR FILING DATE: 2000-08-04
: PRIOR APPLICATION NUMBER: 60/223,378
: PRIOR FILING DATE: 2000-08-07
: NUMBER OF SEQ ID NOS: 9597
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 8927
: LENGTH: 605
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-796-692-8927

Alignment Scores:
Pred. No.: 1.03 Length: 605
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.91% Indels: 0
DB: 9 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-796-692-8927 (1-605)
Oy 940 CysValHisArgAspLeuAlaAlaArgasn 949
|||||
Db 418 TGTGTTCAAGAGACCTGCGCCGACGAGAAC 447

RESULT 53
US-09-796-692-8132/c
: Sequence 8132, Application US/09796692
: Publication No. US20020198362A1
: GENERAL INFORMATION:
: APPLICANT: Gaiger, Alexander
: APPLICANT: Aigate, Paul A.
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
: FILE REFERENCE: 2077.001200
: CURRENT APPLICATION NUMBER: US/09/796,692
: CURRENT FILING DATE: 2001-03-01
: PRIOR APPLICATION NUMBER: 60/186,126
: PRIOR FILING DATE: 2000-03-01
: PRIOR APPLICATION NUMBER: 60/190,479
: PRIOR FILING DATE: 2000-03-17
: PRIOR APPLICATION NUMBER: 60/200,545
: PRIOR FILING DATE: 2000-04-27
: PRIOR APPLICATION NUMBER: 60/200,303
: PRIOR FILING DATE: 2000-04-28
: PRIOR APPLICATION NUMBER: 60/200,779
: PRIOR FILING DATE: 2000-04-28
: PRIOR APPLICATION NUMBER: 60/200,999
: PRIOR FILING DATE: 2000-05-01
: PRIOR APPLICATION NUMBER: 60/202,084
: PRIOR FILING DATE: 2000-05-04
: PRIOR APPLICATION NUMBER: 60/206,201
: PRIOR FILING DATE: 2000-05-22
: PRIOR APPLICATION NUMBER: 60/218,950
: PRIOR FILING DATE: 2000-07-14
: PRIOR APPLICATION NUMBER: 60/222,903
: PRIOR FILING DATE: 2000-08-03
: PRIOR APPLICATION NUMBER: 60/223,416
: PRIOR FILING DATE: 2000-08-04
: PRIOR APPLICATION NUMBER: 60/223,378
: PRIOR FILING DATE: 2000-08-07
: NUMBER OF SEQ ID NOS: 9597
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 8132
: LENGTH: 608
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-796-692-8132

Alignment Scores:
Pred. No.: 1.04 Length: 608
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.91% Indels: 0
DB: 9 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-796-692-8132 (1-608)
Oy 940 CysValHisArgAspLeuAlaAlaArgasn 949
|||||
Db 420 TGTGTTCAAGAGACCTGCGCCGACGAGAAC 391

RESULT 54
US-09-796-692-8085
: Sequence 8085, Application US/09796692
: Publication No. US20020198362A1
: GENERAL INFORMATION:
: APPLICANT: Gaiger, Alexander
: APPLICANT: Aigate, Paul A.
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THER
: FILE REFERENCE: 2077.001200
: CURRENT APPLICATION NUMBER: US/09/796,692
: CURRENT FILING DATE: 2001-03-01
: PRIOR APPLICATION NUMBER: 60/186,126
: PRIOR FILING DATE: 2000-03-01
: PRIOR APPLICATION NUMBER: 60/190,479
: PRIOR FILING DATE: 2000-03-17
: PRIOR APPLICATION NUMBER: 60/200,545
: PRIOR FILING DATE: 2000-04-27
: PRIOR APPLICATION NUMBER: 60/200,303
: PRIOR FILING DATE: 2000-04-28
: PRIOR APPLICATION NUMBER: 60/200,779
: PRIOR FILING DATE: 2000-04-28
: PRIOR APPLICATION NUMBER: 60/200,999
: PRIOR FILING DATE: 2000-05-01
: PRIOR APPLICATION NUMBER: 60/202,084
: PRIOR FILING DATE: 2000-05-04
: PRIOR APPLICATION NUMBER: 60/206,201
: PRIOR FILING DATE: 2000-05-22
: PRIOR APPLICATION NUMBER: 60/218,950
: PRIOR FILING DATE: 2000-07-14
: PRIOR APPLICATION NUMBER: 60/222,903
: PRIOR FILING DATE: 2000-08-03
: PRIOR APPLICATION NUMBER: 60/223,416
: PRIOR FILING DATE: 2000-08-04
: PRIOR APPLICATION NUMBER: 60/223,378
: PRIOR FILING DATE: 2000-08-07
: NUMBER OF SEQ ID NOS: 9597
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 8132
: LENGTH: 608
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-796-692-8132
```

```
: FILE REFERENCE: 2077.001200
: CURRENT APPLICATION NUMBER: US/09/796,692
: CURRENT FILING DATE: 2001-03-01
: PRIOR APPLICATION NUMBER: 60/186,126
: PRIOR FILING DATE: 2000-03-01
: PRIOR APPLICATION NUMBER: 60/190,479
: PRIOR FILING DATE: 2000-03-17
: PRIOR APPLICATION NUMBER: 60/200,545
: PRIOR FILING DATE: 2000-04-27
: PRIOR APPLICATION NUMBER: 60/200,303
: PRIOR FILING DATE: 2000-04-28
: PRIOR APPLICATION NUMBER: 60/200,779
: PRIOR FILING DATE: 2000-04-28
: PRIOR APPLICATION NUMBER: 60/200,999
: PRIOR FILING DATE: 2000-05-01
: PRIOR APPLICATION NUMBER: 60/202,084
: PRIOR FILING DATE: 2000-05-04
: PRIOR APPLICATION NUMBER: 60/206,201
: PRIOR FILING DATE: 2000-05-22
: PRIOR APPLICATION NUMBER: 60/218,950
: PRIOR FILING DATE: 2000-07-14
: PRIOR APPLICATION NUMBER: 60/222,903
: PRIOR FILING DATE: 2000-08-03
: PRIOR APPLICATION NUMBER: 60/223,416
: PRIOR FILING DATE: 2000-08-04
: PRIOR APPLICATION NUMBER: 60/223,378
: PRIOR FILING DATE: 2000-08-07
: NUMBER OF SEQ ID NOS: 9597
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 8132
: LENGTH: 608
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-796-692-8132

Alignment Scores:
Pred. No.: 1.04 Length: 608
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.91% Indels: 0
DB: 9 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-796-692-8132 (1-608)
Oy 940 CysValHisArgAspLeuAlaAlaArgasn 949
|||||
Db 420 TGTGTTCAAGAGACCTGCGCCGACGAGAAC 391

RESULT 54
US-09-796-692-8085
: Sequence 8085, Application US/09796692
: Publication No. US20020198362A1
: GENERAL INFORMATION:
: APPLICANT: Gaiger, Alexander
: APPLICANT: Aigate, Paul A.
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THER
: FILE REFERENCE: 2077.001200
: CURRENT APPLICATION NUMBER: US/09/796,692
: CURRENT FILING DATE: 2001-03-01
: PRIOR APPLICATION NUMBER: 60/186,126
: PRIOR FILING DATE: 2000-03-01
: PRIOR APPLICATION NUMBER: 60/190,479
: PRIOR FILING DATE: 2000-03-17
: PRIOR APPLICATION NUMBER: 60/200,545
: PRIOR FILING DATE: 2000-04-27
: PRIOR APPLICATION NUMBER: 60/200,303
: PRIOR FILING DATE: 2000-04-28
: PRIOR APPLICATION NUMBER: 60/200,779
: PRIOR FILING DATE: 2000-04-28
: PRIOR APPLICATION NUMBER: 60/200,999
: PRIOR FILING DATE: 2000-05-01
: PRIOR APPLICATION NUMBER: 60/202,084
: PRIOR FILING DATE: 2000-05-04
: PRIOR APPLICATION NUMBER: 60/206,201
: PRIOR FILING DATE: 2000-05-22
: PRIOR APPLICATION NUMBER: 60/218,950
: PRIOR FILING DATE: 2000-07-14
: PRIOR APPLICATION NUMBER: 60/222,903
: PRIOR FILING DATE: 2000-08-03
: PRIOR APPLICATION NUMBER: 60/223,416
: PRIOR FILING DATE: 2000-08-04
: PRIOR APPLICATION NUMBER: 60/223,378
: PRIOR FILING DATE: 2000-08-07
: NUMBER OF SEQ ID NOS: 9597
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 8132
: LENGTH: 608
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-796-692-8132
```

```
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO: 8085
; LENGTH: 610
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-796-692-8085

Alignment Scores:
Pred. No.: 1.04 Length: 610
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.91% Indels: 0
DB: 9 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-796-692-8085 (1-610)
QY 940 CysValHisArgAspLeuAlaIaArgAsn 949
DB 418 TGTGTCACAGAGACTGCGCCGACAGAAC 447

RESULT 55
US-09-796-692-7538
; Sequence 7538, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Galiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; FILE REFERENCE: 2077, 001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
```

```
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO: 7538
; LENGTH: 645
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-796-692-7538

Alignment Scores:
Pred. No.: 1.09 Length: 645
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.91% Indels: 0
DB: 9 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-796-692-7538 (1-645)
QY 940 CysValHisArgAspLeuAlaIaArgAsn 949
DB 418 TGTGTCACAGAGACTGCGCCGACAGAAC 447

RESULT 56
US-10-101-464A-368
; Sequence 368, Application US/10101464A
; Publication No. US20030046728A1
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; APPLICANT: Higgins, Colleen M.
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; FILE REFERENCE: 11000,1020C2
; CURRENT APPLICATION NUMBER: US/10/101,464A
; CURRENT FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/704,302
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/228,986
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1998-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 989
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 368
; LENGTH: 869
; TYPE: DNA
; ORGANISM: Pinus radiata
US-10-101-464A-368

Alignment Scores:
Pred. No.: 1.43 Length: 869
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.91% Indels: 0
DB: 9 Gaps: 0

US-09-397-967-16 (1-1099) x US-10-101-464A-368 (1-869)
QY 959 LysIleAlaAspPheGlyLeuAlaLysLeu 968
DB 240 AAGATTGCAGATTGTTGCTGCGCTAAGCTT 269

RESULT 57
US-10-101-464A-218
; Sequence 218, Application US/10101464A
; Publication No. US20030046728A1
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; APPLICANT: Higgins, Colleen M.
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
```

```

; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling
; FILE REFERENCE: 11000.1020c2
; CURRENT APPLICATION NUMBER: US/10/101.464A
; PRIOR FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/704,302
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/228,986
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 989
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 218
; LENGTH: 1021
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-10-101-464A-218

Alignment Scores:
Pred. No.: 1.65      Length: 1021
Score: 10.00      Matches: 10
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 0.91%      Indels: 0
DB: 9      Gaps: 0

US-09-397-967-16 (1-1099) x US-10-101-464A-218 (1-1021)

QY 959 LysilealaaspheglyleuAlaLysleu 968
Db 379 AAGATTGCGATTGTTGGTCTGCAAGCTA 408

RESULT 58
US-10-101-464A-343
; Sequence 343, Application US/10101464A
; Publication No. US20030046728A1
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Neuenhulzen, Nicolaas
; APPLICANT: Higgins, Colleen M.
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; FILE REFERENCE: 11000.1020c2
; CURRENT APPLICATION NUMBER: US/10/101.464A
; CURRENT FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/704,302
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/228,986
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 989
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 343
; LENGTH: 1128
; TYPE: DNA
; ORGANISM: Pinus radiata
US-10-101-464A-343

Alignment Scores:
Pred. No.: 1.8      Length: 1128
Score: 10.00      Matches: 10
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 0.91%      Indels: 0
DB: 9      Gaps: 0

US-09-397-967-16 (1-1099) x US-10-101-464A-343 (1-1128)

```

```

QY 959 LysilealaaspheglyleuAlaLysleu 968
Db 728 AAATAGCAGATTGTTGTTAGCCAACTC 757

RESULT 59
US-10-101-464A-872
; Sequence 872, Application US/10101464A
; Publication No. US20030046728A1
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Neuenhulzen, Nicolaas
; APPLICANT: Higgins, Colleen M.
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; FILE REFERENCE: 11000.1020c2
; CURRENT APPLICATION NUMBER: US/10/101.464A
; CURRENT FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/704,302
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/228,986
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 989
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 872
; LENGTH: 2040
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-10-101-464A-872

Alignment Scores:
Pred. No.: 3.07      Length: 2040
Score: 10.00      Matches: 10
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 0.91%      Indels: 0
DB: 9      Gaps: 0

US-09-397-967-16 (1-1099) x US-10-101-464A-872 (1-2040)

QY 959 LysilealaaspheglyleuAlaLysleu 968
Db 1361 AAGATTGCGATTGTTGGTCTGCAAGCTA 1390

RESULT 60
US-09-938-842A-1497
; Sequence 1497, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SCRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1497
; LENGTH: 2052
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-1497

```

Alignment Scores:

Pred. No.:	3 08	Length:	2052
Score:	10.00	Matches:	10
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	0.91%	Indels:	0
DB:	9	Gaps:	0

US-09-397-967-16 (1-1099) x US-09-938-842A-1497 (1-2052)

OY 959 Lys11eAlaaspheglyLeuAlaLysLeu 968
DB 1465 AAATCGCAGATTTCGATTAGCTAAACTC 1494

RESULT 61

US-09-938-842A-1480
Sequence 1480, Application US/09938842A
Patent No. US20020160378A1
GENERAL INFORMATION:
APPLICANT: Harper, Jeff
APPLICANT: Kreps, Joel
APPLICANT: Wang, Xun
APPLICANT: Zhu, Tong
TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
FILE REFERENCE: SCRIPI300-3
CURRENT APPLICATION NUMBER: US/09/938,842A
CURRENT FILING DATE: 2001-08-24
PRIOR APPLICATION NUMBER: US 60/227,866
PRIOR FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: US 60/264,647
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/300,111
PRIOR FILING DATE: 2001-06-22
NUMBER OF SEQ ID NOS: 5379
SEQ ID NO 1480
LENGTH: 2136
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-938-842A-1480

Alignment Scores:

Pred. No.:	3 2	Length:	2136
Score:	10.00	Matches:	10
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	0.91%	Indels:	0
DB:	9	Gaps:	0

US-09-397-967-16 (1-1099) x US-09-938-842A-1480 (1-2136)

OY 959 Lys11eAlaaspheglyLeuAlaLysLeu 968
DB 1585 AAATCGCAGATTTCGATTAGCTAAACTC 1614

RESULT 62

US-09-924-859A-4
Sequence 4, Application US/09924859A
Patent No. US20020137113A1
GENERAL INFORMATION:
APPLICANT: Godowski, Paul J.
APPLICANT: Mark, Melanie R.
APPLICANT: Sadick, Michael D.
APPLICANT: Shelton, David L.
APPLICANT: Wong, Wai Lee Tan
TITLE OF INVENTION: KINASE RECEPTOR ACTIVATION ASSAY
FILE REFERENCE: P0854C1P2C1
CURRENT APPLICATION NUMBER: US/09/924,859A
CURRENT FILING DATE: 2001-08-08
PRIOR APPLICATION NUMBER: US/09/417,381
PRIOR FILING DATE: 1999-10-13
NUMBER OF SEQ ID NOS: 11

SEQ ID NO 4
LENGTH: 2820
TYPE: DNA
ORGANISM: Homo Sapien
US-09-924-859A-4

Alignment Scores:

Pred. No.:	4 1	Length:	2820
Score:	10.00	Matches:	10
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	0.91%	Indels:	0
DB:	10	Gaps:	0

US-09-397-967-16 (1-1099) x US-09-924-859A-4 (1-2820)

OY 1003 SeraspValTrpSerPheglyValValLeu 1012
DB 2468 AGCGACGTCTGGAGCTTCGCGCTGCTC 2497

RESULT 63

US-09-982-610-22
Sequence 22, Application US/09982610
Patent No. US20020146420A1
GENERAL INFORMATION:
APPLICANT: Genentech, Inc.
Bennett, Brian D.
Goeddel, David
Lee, James M.
Mathews, William
Tsai, Siao Ping
Wood, William I.
TITLE OF INVENTION: PROTEIN TYROSINE KINASE ACONIST ANTIBODIES
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESS: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Winpatin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/982,610
FILING DATE: 17-Oct-2001
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/446,648
FILING DATE: 1996-MAY-23
APPLICATION NUMBER: 08/222616
FILING DATE: 04-APR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M.
REGISTRATION NUMBER: 40,378
REFERENCE/DOCKET NUMBER: P0821P3PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1994
TELEFAX: 415/952-9881
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 3120 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Single
TOPOLOGY: Linear
SEQUENCE DESCRIPTION: SEQ ID NO: 22:
US-09-982-610-22

Alignment Scores:

```
Pred. No.: 4.49 Length: 3120
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.91% Indels: 0
DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-982-610-22 (1-3120)

QY 940 CysValHisArgAspLeuAlaIaArgAsn 949
|||||
Db 2419 TGTGTCCACGACGCTGCGCCGACGAC 2448

RESULT 64
US-09-866-510-1
; Sequence 1, Application US/09866510
; Patent No. US20020111304A1
; GENERAL INFORMATION:
; APPLICANT: KAZLAUSKAS, ANDRIUS
; APPLICANT: IKUNO, YASUSHI
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING FIBROTIC DISEASES
; FILE REFERENCE: ERM-104.01
; CURRENT APPLICATION NUMBER: US/09/866,510
; CURRENT FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/250,747
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: 60/289,103
; PRIOR FILING DATE: 2001-05-07
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 3270
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3267)
US-09-866-510-1

Alignment Scores:
Pred. No.: 4.68 Length: 3270
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.91% Indels: 0
DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-866-510-1 (1-3270)

QY 940 CysValHisArgAspLeuAlaIaArgAsn 949
|||||
Db 2440 TGTGTCCACGCGTGTGCTGCTGCAAC 2469

RESULT 65
US-09-866-510-3
; Sequence 3, Application US/09866510
; Patent No. US20020111304A1
; GENERAL INFORMATION:
; APPLICANT: KAZLAUSKAS, ANDRIUS
; APPLICANT: IKUNO, YASUSHI
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING FIBROTIC DISEASES
; FILE REFERENCE: ERM-104.01
; CURRENT APPLICATION NUMBER: US/09/866,510
; CURRENT FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/250,747
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: 60/289,103
; PRIOR FILING DATE: 2001-05-07
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 3270
; TYPE: DNA
```

```
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3267)
US-09-866-510-3

Alignment Scores:
Pred. No.: 4.68 Length: 3270
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.91% Indels: 0
DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-866-510-3 (1-3270)

QY 940 CysValHisArgAspLeuAlaIaArgAsn 949
|||||
Db 2440 TGTGTCCACGCGTGTGCTGCTGCAAC 2469

RESULT 66
US-09-866-510-5
; Sequence 5, Application US/09866510
; Patent No. US20020111304A1
; GENERAL INFORMATION:
; APPLICANT: KAZLAUSKAS, ANDRIUS
; APPLICANT: IKUNO, YASUSHI
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING FIBROTIC DISEASES
; FILE REFERENCE: ERM-104.01
; CURRENT APPLICATION NUMBER: US/09/866,510
; CURRENT FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/250,747
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: 60/289,103
; PRIOR FILING DATE: 2001-05-07
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 3270
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3267)
US-09-866-510-5

Alignment Scores:
Pred. No.: 4.68 Length: 3270
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.91% Indels: 0
DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-866-510-5 (1-3270)

QY 940 CysValHisArgAspLeuAlaIaArgAsn 949
|||||
Db 2440 TGTGTCCACGCGTGTGCTGCTGCAAC 2469

RESULT 67
US-09-866-510-9
; Sequence 9, Application US/09866510
; Patent No. US20020111304A1
; GENERAL INFORMATION:
; APPLICANT: KAZLAUSKAS, ANDRIUS
; APPLICANT: IKUNO, YASUSHI
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING FIBROTIC DISEASES
; FILE REFERENCE: ERM-104.01
; CURRENT APPLICATION NUMBER: US/09/866,510
; CURRENT FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/250,747
; PRIOR FILING DATE: 2000-12-01
```



```

; PRIOR APPLICATION NUMBER: 60/289,103
; PRIOR FILING DATE: 2001-05-07
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 3270
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3267)
US-09-866-510-9

Alignment Scores:
Pred. No.: 4.68 Length: 3270
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.91% Indels: 0
DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-866-510-9 (1-3270)
QY 940 CysValHisArgAspLeuAlaAlaArgAsn 949
Db 2440 TGTGTCACACGCTGATCGCTGCTGCAGAC 2469

RESULT 68
US-09-866-510-13
; Sequence 13, Application US/09866510
; Patent No. US20020111304A1
; GENERAL INFORMATION:
; APPLICANT: KAZLAUSKAS, ANDRIUS
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING FIBROTIC DISEASES
; FILE REFERENCE: ERM-104.01
; CURRENT APPLICATION NUMBER: US/09/866,510
; CURRENT FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/250,747
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: 60/289,103
; PRIOR FILING DATE: 2001-05-07
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 3321
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3270)
US-09-866-510-13

Alignment Scores:
Pred. No.: 4.74 Length: 3321
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.91% Indels: 0
DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-866-510-13 (1-3321)
QY 940 CysValHisArgAspLeuAlaAlaArgAsn 949
Db 2464 TCGGTCACACGAGACTGCGGCTAGAGAC 2493

RESULT 69
US-09-866-510-15
; Sequence 15, Application US/09866510
; Patent No. US20020111304A1
; GENERAL INFORMATION:
; APPLICANT: KAZLAUSKAS, ANDRIUS
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING FIBROTIC DISEASES
; FILE REFERENCE: ERM-104.01
; CURRENT APPLICATION NUMBER: US/09/866,510
; CURRENT FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/250,747
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: 60/289,103
; PRIOR FILING DATE: 2001-05-07
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 3321
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3318)
US-09-866-510-15

Alignment Scores:
Pred. No.: 4.74 Length: 3321
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.91% Indels: 0
DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-866-510-15 (1-3321)
QY 940 CysValHisArgAspLeuAlaAlaArgAsn 949
Db 2464 TCGGTCACACGAGACTGCGGCTAGAGAC 2493

RESULT 70
US-09-866-510-17
; Sequence 17, Application US/09866510
; Patent No. US20020111304A1
; GENERAL INFORMATION:
; APPLICANT: KAZLAUSKAS, ANDRIUS
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING FIBROTIC DISEASES
; FILE REFERENCE: ERM-104.01
; CURRENT APPLICATION NUMBER: US/09/866,510
; CURRENT FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/250,747
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: 60/289,103
; PRIOR FILING DATE: 2001-05-07
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
; LENGTH: 3321
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3318)
US-09-866-510-17

Alignment Scores:
Pred. No.: 4.74 Length: 3321
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.91% Indels: 0
DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-866-510-17 (1-3321)
QY 940 CysValHisArgAspLeuAlaAlaArgAsn 949
Db 2464 TCGGTCACACGAGACTGCGGCTAGAGAC 2493
```

RESULT 71
US-09-866-510-21
; Sequence 21, Application US/09866510
; Patent No. US20020111304A1
; GENERAL INFORMATION:
; APPLICANT: KAZLAUSKAS, ANDRIUS
; APPLICANT: IKONO, YASUSHI
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING FIBROTIC DISEASES
; FILE REFERENCE: ERM-104.01
; CURRENT APPLICATION NUMBER: 2001-05-25
; PRIORITY FILING DATE: 2001-05-25
; PRIORITY FILING DATE: 2000-12-01
; PRIORITY FILING DATE: 2000-12-01
; PRIORITY FILING DATE: 2001-05-07
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 3321
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3318)
US-09-866-510-21

Alignment Scores:
Pred. No.: 4.74 Length: 3321
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.91% Indels: 0
Gaps: 0

US-09-397-967-16 (1-1099) x US-09-866-510-21 (1-3321)

QY 940 CysValHisArgAspLeuAlaIaArgAsn 949
|||||
DB 2464 TGCCTCCACAGACCTGGCGCTAGGAAAC 2493

RESULT 72
US-09-919-408-1
; Sequence 1, Application US/09919408
; Patent No. US20020072077A1
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ihor R.
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
; RECEPTORS AND THEIR LIGANDS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Imclone Systems Incorporated
; STREET: 180 Varick Street
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10014
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/919,408
; FILING DATE: 31-Jul-2001
; CLASSIFICATION: <Unknown>
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 07/977,451
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 07/906,397
; FILING DATE: 26-JUN-1992
; APPLICATION NUMBER: US PCT/US92/05401
; FILING DATE: 26-JUN-1992

APPLICATION NUMBER: TW 81102961
FILING DATE: 15-APR-1992
APPLICATION NUMBER: US PCT/US92/02750
FILING DATE: 02-APR-1992
APPLICATION NUMBER: US 07/813,593
FILING DATE: 24-DEC-1991
APPLICATION NUMBER: US 07/793,065
FILING DATE: 15-NOV-1991
APPLICATION NUMBER: US 07/728,913
FILING DATE: 28-JUN-1991
APPLICATION NUMBER: US 07/679,666
FILING DATE: 02-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Feil, Irving N.
REGISTRATION NUMBER: 28,601
REFERENCE/DOCKET NUMBER: LEM-3-7P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-645-1405
TELEFAX: 212-645-2054
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3453 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 112..3006
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 31..111
FEATURE:
NAME/KEY: CDS
LOCATION: 31..3009
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-919-408-1

Alignment Scores:
Pred. No.: 4.91 Length: 3453
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.91% Indels: 0
Gaps: 0

US-09-397-967-16 (1-1099) x US-09-919-408-1 (1-3453)

QY 940 CysValHisArgAspLeuAlaIaArgAsn 949
|||||
DB 2458 TGTGTCCACAGACCTGGCAGCCAGGAT 2487

RESULT 73
US-09-872-136-1
; Sequence 1, Application US/09872136
; Patent No. US20020119545A1
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ihor R.
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
; RECEPTORS AND THEIR LIGANDS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Imclone Systems Incorporated
; STREET: 180 Varick Street
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10014
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/872,136
FILING DATE: 01-Jun-2001
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/208,786
FILING DATE: <unknown>
APPLICATION NUMBER: US/09/021,324
FILING DATE: <unknown>
APPLICATION NUMBER: US/07/977,451
FILING DATE: 1992-11-19
APPLICATION NUMBER: US 07/906,397
FILING DATE: 26-JUN-1992
APPLICATION NUMBER: US PCT/US92/05401
FILING DATE: 26-JUN-1992
APPLICATION NUMBER: TW 81102961
FILING DATE: 15-APR-1992
APPLICATION NUMBER: US PCT/US92/02750
FILING DATE: 02-APR-1992
APPLICATION NUMBER: US 07/813,593
FILING DATE: 24-DEC-1991
APPLICATION NUMBER: US 07/793,065
FILING DATE: 15-NOV-1991
APPLICATION NUMBER: US 07/728,913
FILING DATE: 28-JUN-1991
APPLICATION NUMBER: US 07/679,666
FILING DATE: 02-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Felt, Irving N.
REGISTRATION NUMBER: 28,601
REFERENCE/DOCKET NUMBER: LEM-3-7P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-645-1405
TELEFAX: 212-645-2054
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3453 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 112..3006
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 31..111
FEATURE:
NAME/KEY: CDS
LOCATION: 31..3009
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-872-136-1
Alignment Scores:
Pred. No.: 4.91 Length: 3453
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.91% Indels: 0
DB: 10 Gaps: 0
US-09-397-967-16 (1-1099) x US-09-872-136-1 (1-3453)
QY 940 CysValHisArgAspLeuAlaIaArgAsn 949
|||||
DB 2458 TGTGTCACAGAGACCTGCGCAGCGAAT 2487

RESULT 74
US-09-919-408-3
Sequence 3, Application US/09919408
Patent No. US20020072077A1
GENERAL INFORMATION:
APPLICANT: Lemischka, Ihor R.
TITLE OF INVENTION: TORIPOTENT HEMATOPOIETIC STEM CELL
RECEPTORS AND THEIR LIGANDS
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Imclone Systems Incorporated
STREET: 180 Varick Street
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10014
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/919,408
FILING DATE: 31-Jul-2001
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/977,451
FILING DATE: <unknown>
APPLICATION NUMBER: US 07/906,397
FILING DATE: 26-JUN-1992
APPLICATION NUMBER: US PCT/US92/05401
FILING DATE: 26-JUN-1992
APPLICATION NUMBER: TW 81102961
FILING DATE: 15-APR-1992
APPLICATION NUMBER: US PCT/US92/02750
FILING DATE: 02-APR-1992
APPLICATION NUMBER: US 07/813,593
FILING DATE: 24-DEC-1991
APPLICATION NUMBER: US 07/793,065
FILING DATE: 15-NOV-1991
APPLICATION NUMBER: US 07/728,913
FILING DATE: 28-JUN-1991
APPLICATION NUMBER: US 07/679,666
FILING DATE: 02-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Felt, Irving N.
REGISTRATION NUMBER: 28,601
REFERENCE/DOCKET NUMBER: LEM-3-7P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-645-1405
TELEFAX: 212-645-2054
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3501 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
FEATURE:
NAME/KEY: CDS
LOCATION: 58..3039
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 139..3036
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 58..138
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-919-408-3

Alignment Scores:
Pred. No.: 4.97 Length: 3501
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.91% Indels: 0
DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-919-408-3 (1-3501)

QY 940 CysvalHisArgAspLeuAlaIArgAsn 949
|||||
Db 2476 TGTGTTCACAGACCTGGCCGACGAGAAC 2505

RESULT 75
US-09-872-136-3
Sequence 3, Application US/09872136
Patent No. US20020119545A1

GENERAL INFORMATION:
APPLICANT: Lemischka, Ihor R.
TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
RECEPTORS AND THEIR LIGANDS
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Imclone Systems Incorporated
STREET: 180 Varlick Street
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10014

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/872,136
FILING DATE: 01-Jun-2001
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/208,786
FILING DATE: <Unknown>
APPLICATION NUMBER: US/09/021,324
FILING DATE: <Unknown>
APPLICATION NUMBER: US/07/977,451
FILING DATE: 1992-11-19
APPLICATION NUMBER: US 07/906,397
FILING DATE: 26-JUN-1992
APPLICATION NUMBER: US PCT/US92/05401
FILING DATE: 26-JUN-1992
APPLICATION NUMBER: TW 81102961
FILING DATE: 15-APR-1992
APPLICATION NUMBER: US PCT/US92/02750
FILING DATE: 02-APR-1992
APPLICATION NUMBER: US 07/813,593
FILING DATE: 24-DEC-1991
APPLICATION NUMBER: US 07/793,065
FILING DATE: 15-NOV-1991
APPLICATION NUMBER: US 07/728,913
FILING DATE: 28-JUN-1991
APPLICATION NUMBER: US 07/679,666
FILING DATE: 02-APR-1991

ATTORNEY/AGENT INFORMATION:
NAME: Felt, Irving N.
REGISTRATION NUMBER: 28,601
REFERENCE/DOCKET NUMBER: LEM-3-7P

TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-645-1405
TELEFAX: 212-645-2054

INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3501 base pairs
TYPE: nucleic acid

STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal

FEATURE:
NAME/KEY: CDS
LOCATION: 58..3039

FEATURE:
NAME/KEY: mat_peptide
LOCATION: 139..3036

FEATURE:
NAME/KEY: sig_peptide
LOCATION: 58..138

SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-872-136-3

Alignment Scores:
Pred. No.: 4.97 Length: 3501
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.91% Indels: 0
DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-872-136-3 (1-3501)

QY 940 CysvalHisArgAspLeuAlaIArgAsn 949
|||||
Db 2476 TGTGTTCACAGACCTGGCCGACGAGAAC 2505

RESULT 76
US-09-955-363-35
Sequence 35, Application US/09955363
Patent No. US20020173621A1

GENERAL INFORMATION:
APPLICANT: Sledziewski Ph.D., Andrzej Z
Bell, Lilijan A.
Kindsvogel Ph.D., Wayne R.
TITLE OF INVENTION: METHODS OF PRODUCING SECRETED RECEPTOR ANALOGS
AND BIOLOGICALLY ACTIVE DIMERIZED POLYPEPTIDE
FUSIONS

NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Seed and Berry
STREET: 6300 Columbia Center
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98104-7092

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.24

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/955,363
FILING DATE: 18-Sep-2001
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/634,510
FILING DATE: <Unknown>
APPLICATION NUMBER: US 07/347,291
FILING DATE: 02-MAY-1989

ATTORNEY/AGENT INFORMATION:
NAME: Makl J.D., David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 990008.446C3

TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-622-4900
TELEFAX: 206-682-6031
TELEX: 3723836

```

; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 4054 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: linear
;   MOLECULE TYPE: cDNA
;   HYPOTHETICAL: N
;   ANTI-SENSE: N
;   ORIGINAL SOURCE:
;     ORGANISM: Homo sapiens
;     IMMEDIATE SOURCE:
;     CLONE: p-alpha-17B
;   FEATURE:
;     NAME/KEY: CDS
;     LOCATION: 205..3471
;   OTHER INFORMATION:
;
; SEQUENCE DESCRIPTION: SEQ ID NO: 35:
US-09-955-363-35

Alignment Scores:
Pred. No.: 5.67 Length: 4054
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.91% Indels: 0
DB: 9 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-955-363-35 (1-4054)
QY 940 CysvalHisArgAspLeuAlaAlaArgAsn 949
Db 2644 TGATGTCACCGATGCTGCTGCTGCCAAC 2673

RESULT 77
US-10-022-939-1
; Sequence 1, Application US/10022939
; Publication No. US20030032160A1
; GENERAL INFORMATION:
;   APPLICANT: Kendall, Richard L.
;   APPLICANT: Thomas, Kenneth A.
;   APPLICANT: Mao, Xianzhi
;   APPLICANT: Tedden, Andrew
;   TITLE OF INVENTION: HUMAN RECEPTOR TYROSINE KINASE, KDR
;   FILE REFERENCE: 19963YDB
;   CURRENT APPLICATION NUMBER: US/10/022,939
;   CURRENT FILING DATE: 2001-12-18
;   PRIOR APPLICATION NUMBER: 09/483,539
;   PRIOR FILING DATE: 2000-01-14
;   PRIOR APPLICATION NUMBER: 09/098,707
;   PRIOR FILING DATE: 1998-06-17
;   PRIOR APPLICATION NUMBER: 60/050,962
;   PRIOR FILING DATE: 1997-06-18
;   NUMBER OF SEQ ID NOS: 8
;   SOFTWARE: FastSeq for Windows Version 4.0
;   SEQ ID NO 1
;   LENGTH: 4071
;   TYPE: DNA
;   ORGANISM: Human
;   US-10-022-939-1

Alignment Scores:
Pred. No.: 5.69 Length: 4071
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.91% Indels: 0
DB: 9 Gaps: 0

US-09-397-967-16 (1-1099) x US-10-022-939-1 (1-4071)
QY 942 HisArgAspLeuAlaAlaArgAsnIleLeu 951
Db 942 HisArgAspLeuAlaAlaArgAsnIleLeu 951

US-09-375-248-1
; Sequence 1, Application US/09375248
; Publication No. US20030026759A1
; GENERAL INFORMATION:
;   APPLICANT: Fretell, Robert E.
;   APPLICANT: Altalo, Karl
;   APPLICANT: Finegold, David N.
;   APPLICANT: Karkkainen, Marika
;   TITLE OF INVENTION: SCREENING AND THERAPY FOR LYMPHATIC DISORDERS INVOLVING
;   FILE REFERENCE: 28967/35255A
;   CURRENT APPLICATION NUMBER: US/09/375,248
;   CURRENT FILING DATE: 1999-08-16
;   EARLIER APPLICATION NUMBER: PCT/US99/06133
;   EARLIER FILING DATE: 1999-03-26
;   NUMBER OF SEQ ID NOS: 28
;   SOFTWARE: PatentIn Ver. 2.0
;   SEQ ID NO 1
;   LENGTH: 4111
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;     NAME/KEY: CDS
;     LOCATION: (20)..(4111)
;   OTHER INFORMATION: Human Flt4 (VEGFR-3) long form cDNA
;   US-09-375-248-1

US-10-100-405A-1
; Sequence 1, Application US/10100405A
; Publication No. US20030055239A1
; GENERAL INFORMATION:
;   APPLICANT: Kendall, Richard L.
;   APPLICANT: Thomas, Kenneth A.
;   APPLICANT: Mao, Xianzhi
;   APPLICANT: Tedden, Andrew
;   TITLE OF INVENTION: HUMAN RECEPTOR TYROSINE KINASE, KDR
;   FILE REFERENCE: 19963YDC
;   CURRENT APPLICATION NUMBER: US/10/100,405A
;   CURRENT FILING DATE: 2002-08-13
;   PRIOR APPLICATION NUMBER: 10/022,939
;   PRIOR FILING DATE: 2001-12-18
;   PRIOR APPLICATION NUMBER: 09/483,539
;   PRIOR FILING DATE: 2000-01-14
;   PRIOR APPLICATION NUMBER: 09/098,707
;   PRIOR FILING DATE: 1998-06-17
;   PRIOR APPLICATION NUMBER: 60/050,962
;   PRIOR FILING DATE: 1997-06-18
;   NUMBER OF SEQ ID NOS: 8
;   SOFTWARE: FastSeq for Windows Version 4.0
;   SEQ ID NO 1
;   LENGTH: 4071
;   TYPE: DNA
;   ORGANISM: Human
;   US-10-100-405A-1

Alignment Scores:
Pred. No.: 5.69 Length: 4071
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.91% Indels: 0
DB: 9 Gaps: 0

US-09-397-967-16 (1-1099) x US-10-100-405A-1 (1-4071)
QY 942 HisArgAspLeuAlaAlaArgAsnIleLeu 951
Db 3076 CACAGGACCTGGCGGCACGAAATATTCCTC 3105

RESULT 79
US-09-375-248-1
; Sequence 1, Application US/09375248
; Publication No. US20030026759A1
; GENERAL INFORMATION:
;   APPLICANT: Fretell, Robert E.
;   APPLICANT: Altalo, Karl
;   APPLICANT: Finegold, David N.
;   APPLICANT: Karkkainen, Marika
;   TITLE OF INVENTION: SCREENING AND THERAPY FOR LYMPHATIC DISORDERS INVOLVING
;   FILE REFERENCE: 28967/35255A
;   CURRENT APPLICATION NUMBER: US/09/375,248
;   CURRENT FILING DATE: 1999-08-16
;   EARLIER APPLICATION NUMBER: PCT/US99/06133
;   EARLIER FILING DATE: 1999-03-26
;   NUMBER OF SEQ ID NOS: 28
;   SOFTWARE: PatentIn Ver. 2.0
;   SEQ ID NO 1
;   LENGTH: 4111
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;     NAME/KEY: CDS
;     LOCATION: (20)..(4111)
;   OTHER INFORMATION: Human Flt4 (VEGFR-3) long form cDNA
;   US-09-375-248-1
```

Alignment Scores:

Pred. No.: 5 74 Length: 4111
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.91% Indels: 0
DB: 9 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-375-248-1 (1-4111)

OY 942 HisargaspLeuAlaAlaArgAsnIleLeu 951

Db 3122 CACAGAGACGTGCTGCTCGACATCTCTG 3151

RESULT 80

US-09-982-610-31

Sequence 31, Application US/09982610

Patent No. US20020146420A1

GENERAL INFORMATION:

APPLICANT: Genentech, Inc.

Bennett, Brian D.

Goeddel, David

Lee, James M.

Matthews, William

Tsai, Siao Ping

Wood, William I.

ADDRESS: Genentech, Inc.

STREET: 460 Point San Bruno Blvd

CITY: South San Francisco

STATE: California

COUNTRY: USA

ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Winpatin (Genentech)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/982,610

FILING DATE: 17-Oct-2001

CLASSIFICATION: <unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/446,648

FILING DATE: 1996-MAY-23

APPLICATION NUMBER: 08/222616

FILING DATE: 04-APR-1994

ATTORNEY/AGENT INFORMATION:

NAME: Lee, Wendy M.

REGISTRATION NUMBER: 40,378

REFERENCE/DOCKET NUMBER: P0821P3PCT

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415/225-1994

TELEFAX: 415/952-9881

TELEX: 910/371-7168

INFORMATION FOR SEQ ID NO: 31:

LENGTH: 4425 base pairs

TYPE: Nucleic Acid

STRANDEDNESS: Single

TOPOLOGY: Linear

SEQUENCE DESCRIPTION: SEQ ID NO: 31:

US-09-982-610-31

Alignment Scores:

Pred. No.: 6 13 Length: 4425
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.91% Indels: 0

DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-982-610-31 (1-4425)

OY 942 HisargaspLeuAlaAlaArgAsnIleLeu 951

Db 3133 CACAGAGACGTGCTGCTCGACATCTCTG 3162

RESULT 81

US-09-955-363-1

Sequence 1, Application US/09955363

Patent No. US20020173621A1

GENERAL INFORMATION:

APPLICANT: Sledziewski Ph.D., Andrzej Z

Bell, Lillian A.

Kindvogel Ph.D., Wayne R.

TITLE OF INVENTION: METHODS OF PRODUCING SECRETED RECEPTOR ANALOGS

AND BIOLOGICALLY ACTIVE DIMERIZED POLYPEPTIDE

FUSIONS

NUMBER OF SEQUENCES: 36

CORRESPONDENCE ADDRESS:

ADDRESS: Seed and Berry

STREET: 6300 Columbia Center

CITY: Seattle

STATE: WA

COUNTRY: USA

ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.24

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/955,363

FILING DATE: 18-Sep-2001

CLASSIFICATION: <unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/634,510

FILING DATE: <unknown>

APPLICATION NUMBER: US 07/347,291

FILING DATE: 02-MAY-1989

ATTORNEY/AGENT INFORMATION:

NAME: Makl J.D., David J.

REGISTRATION NUMBER: 31,392

REFERENCE/DOCKET NUMBER: 990008.446C3

TELECOMMUNICATION INFORMATION:

TELEPHONE: 206-622-4900

TELEFAX: 206-682-6031

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 4465 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: CDNA

HYPOTHETICAL: N

ANTI-SENSE: N

ORIGINAL SOURCE:

ORGANISM: Homo sapiens

DEVELOPMENTAL STAGE: Adult

TISSUE TYPE: Skin

CELL TYPE: fibroblasts

IMMEDIATE SOURCE:

CLONE: PR-1X1

FEATURE:

NAME/KEY: CDS

LOCATION: 354..3671

OTHER INFORMATION:

SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-09-955-363-1

Alignment Scores:

Pred. No.: 354..3671 Length: 4425
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.91% Indels: 0

```

Pred. No.: 6.18 Length: 4465
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.91% Indels: 0
DB: 9 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-955-363-1 (1-4465)

QY 940 CysValHisArgAspLeuAlaAlaArgAsn 949
DB 2817 TGGCTCCACAGAGACCTGGCGGCTAGGAC 2846

RESULT 82
US-09-870-759-119
; Sequence 119, Application US/09870759
; Patent No. US20020177551A1
; GENERAL INFORMATION:
; APPLICANT: TERMAN, David S
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
; FILE REFERENCE: 870759
; CURRENT APPLICATION NUMBER: US/09/870,759
; PRIOR FILING DATE: 2002-01-14
; PRIOR APPLICATION NUMBER: US 60/208,128
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 166
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 119
; LENGTH: 4989
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (46)..(4149)
; OTHER INFORMATION:
US-09-870-759-119

Alignment Scores:
Pred. No.: 6.83 Length: 4989
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.91% Indels: 0
DB: 9 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-870-759-119 (1-4989)

QY 1003 SerAspValTrpSerPheGlyValValLeu 1012
DB 3622 TCGGACGCTGTGCTTCGCGGCTGCTC 3651

RESULT 83
US-09-954-531-1383
; Sequence 1383, Application US/09954531
; Patent No. US20020165180A1
; GENERAL INFORMATION:
; APPLICANT: Weaver, Zoe
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents using Canc
; FILE REFERENCE: 689290-77
; CURRENT APPLICATION NUMBER: US/09/954,531
; PRIOR FILING DATE: 2002-05-02
; PRIOR APPLICATION NUMBER: US/60/233,133
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,009
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,034
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,509
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: US/60/234,567
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 1392
```

```

; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1383
; LENGTH: 5084
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-531-1383

Alignment Scores:
Pred. No.: 6.94 Length: 5084
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.91% Indels: 0
DB: 9 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-954-531-1383 (1-5084)

QY 942 HisArgAspLeuAlaAlaArgAsnIleLeu 951
DB 2389 CACAGAGACTTGGCAGCCAGAAATATCTC 2418

RESULT 84
US-09-967-768A-277
; Sequence 277, Application US/09967768A
; Patent No. US20020150877A1
; GENERAL INFORMATION:
; APPLICANT: Augustus, Meena
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Sign
; FILE REFERENCE: 689290-72
; CURRENT APPLICATION NUMBER: US/09/967,768A
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US/60/236,109
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,034
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,111
; PRIOR FILING DATE: 2000-09-28
; NUMBER OF SEQ ID NOS: 325
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 277
; LENGTH: 5084
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-967-768A-277

Alignment Scores:
Pred. No.: 6.94 Length: 5084
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.91% Indels: 0
DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-967-768A-277 (1-5084)

QY 942 HisArgAspLeuAlaAlaArgAsnIleLeu 951
DB 2389 CACAGAGACTTGGCAGCCAGAAATATCTC 2418

RESULT 85
US-09-919-408-5
; Sequence 5, Application US/09919408
; Patent No. US20020072077A1
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Thor R.
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
; RECEPTORS AND THEIR LIGANDS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Imclone Systems Incorporated
; STREET: 180 Varick Street
; CITY: New York
```

STATE: New York
COUNTRY: U.S.A.
ZIP: 10014
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/919,408
FILING DATE: 31-Jul-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/977,451
FILING DATE: <Unknown>
APPLICATION NUMBER: US 07/906,397
FILING DATE: 26-JUN-1992
APPLICATION NUMBER: US PCT/US92/05401
FILING DATE: 26-JUN-1992
APPLICATION NUMBER: TW 81102961
FILING DATE: 15-APR-1992
APPLICATION NUMBER: US PCT/US92/02750
FILING DATE: 02-APR-1992
APPLICATION NUMBER: US 07/813,593
FILING DATE: 24-DEC-1991
APPLICATION NUMBER: US 07/793,065
FILING DATE: 15-NOV-1991
APPLICATION NUMBER: US 07/728,913
FILING DATE: 28-JUN-1991
APPLICATION NUMBER: US 07/679,666
FILING DATE: 02-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Felt, Irving N.
REGISTRATION NUMBER: 28,601
REFERENCE/DOCKET NUMBER: LEW-3-7P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-645-1405
TELEFAX: 212-645-2054
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 5406 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
FEATURE:
NAME/KEY: CDS
LOCATION: 208..4311
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 265..4308
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 208..264
SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-919-408-5
Alignment Scores:
Pred. No.: 7.34 Length: 5406
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.91% Indels: 0
DB: 10 Gaps: 0
US-09-397-967-16 (1-1099) x US-09-919-408-5 (1-5406)
QY 942 HisArgAspLeuAlaAlaArgAsnIleLeu 951
DB 3277 CACAGGAGCCTGGCAGCAGAACATTTTC 3306

RESULT 86
US-09-872-136-5
Sequence 5, Application US/09872136
Patent No. US20020119545A1
GENERAL INFORMATION:
APPLICANT: Lemischka, Ihor R.
TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
RECEPTORS AND THEIR LIGANDS
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Imclone Systems Incorporated
STREET: 180 Varlick Street
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10014
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/872,136
FILING DATE: 01-Jun-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/208,786
FILING DATE: <Unknown>
APPLICATION NUMBER: US/09/021,324
FILING DATE: <Unknown>
APPLICATION NUMBER: US/07/977,451
FILING DATE: 1992-11-19
APPLICATION NUMBER: US 07/906,397
FILING DATE: 26-JUN-1992
APPLICATION NUMBER: US PCT/US92/05401
FILING DATE: 26-JUN-1992
APPLICATION NUMBER: TW 81102961
FILING DATE: 15-APR-1992
APPLICATION NUMBER: US PCT/US92/02750
FILING DATE: 02-APR-1992
APPLICATION NUMBER: US 07/813,593
FILING DATE: 24-DEC-1991
APPLICATION NUMBER: US 07/793,065
FILING DATE: 15-NOV-1991
APPLICATION NUMBER: US 07/728,913
FILING DATE: 28-JUN-1991
APPLICATION NUMBER: US 07/679,666
FILING DATE: 02-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Felt, Irving N.
REGISTRATION NUMBER: 28,601
REFERENCE/DOCKET NUMBER: LEW-3-7P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-645-1405
TELEFAX: 212-645-2054
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 5406 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
FEATURE:
NAME/KEY: CDS
LOCATION: 208..4311
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 265..4308
FEATURE:


```

; NAME/KEY: sig-peptide
; LOCATION: 208..264
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-872-136-5

Alignment Scores:
Pred. No.: 7.34 Length: 5406
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.91% Indels: 0
DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-872-136-5 (1-5406)
Oy 942 HsargaspLeuAlaAlaargasnIleLeu 951
Db 3277 CACAGGACCTGGCAGCAGAAACATTCTC 3306

RESULT 87
US-09-766-678-1
; Sequence 1, Application US/09766678
; Patent No. US20020081650A1
; GENERAL INFORMATION:
; APPLICANT: Ullrich, Axel
; Rissau, Werner
; Millauer, Birgit
; Gazit, Aviv
; Levitzki, Alex
; TITLE OF INVENTION: Flk-1 Is A Receptor For Vascular
; Endothelial Growth Factor
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/766,678
; FILING DATE: 25-Jan-2001
; CLASSIFICATION: <unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/193,829
; FILING DATE: 09-FEB-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7683-060
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)790-9090
; TELEFAX: (212)869-9741
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5470 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 286..4386
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-766-678-1

Alignment Scores:
Pred. No.: 7.41 Length: 5470
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.91% Indels: 0
DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-766-678-1 (1-5470)
Oy 942 HsargaspLeuAlaAlaargasnIleLeu 951
Db 3355 CACAGGACCTGGCAGCAGAAACATTCTC 3384

RESULT 88
US-09-919-497-40
; Sequence 40, Application US/09919497
; Patent No. US20020106662A1
; GENERAL INFORMATION:
; APPLICANT: Mutter, George L.
; TITLE OF INVENTION: PROGNOSTIC CLASSIFICATION OF ENDOMETRIAL CANCER
; FILE REFERENCE: B0801/7225
; CURRENT APPLICATION NUMBER: US/09/919,497
; CURRENT FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 60/221,735
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 40
; LENGTH: 6378
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-919-497-40

Alignment Scores:
Pred. No.: 8.51 Length: 6378
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.91% Indels: 0
DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-919-497-40 (1-6378)
Oy 940 CysvalHsargaspLeuAlaAlaargasn 949
Db 2579 TGTGTCCACCGATCTGGCTGCTCGCAAC 2608

RESULT 89
US-09-769-987-1
; Sequence 1, Application US/09769987
; Patent No. US20020055129A1
; GENERAL INFORMATION:
; APPLICANT: Matsui, Toshimitsu
; APPLICANT: Aaronson, Stuart A.
; APPLICANT: Pierce, Jacalyn H.
; TITLE OF INVENTION: Antibodies for the Alpha Platelet-Derived Growth Factor Recept
; FILE REFERENCE: 14014.026602
; CURRENT APPLICATION NUMBER: US/09/769,987
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 08/460,656
; PRIOR FILING DATE: 1995-06-02
; PRIOR APPLICATION NUMBER: US 08/439,095
; PRIOR FILING DATE: 1995-05-11
; PRIOR APPLICATION NUMBER: US 07/915,884
; PRIOR FILING DATE: 1992-07-20
; PRIOR APPLICATION NUMBER: US 07/308,282
; PRIOR FILING DATE: 1989-02-09
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 6412
; TYPE: DNA
; ORGANISM: Artificial Sequence
```

FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: No. US2002005129a1e =
OTHER INFORMATION: synthetic construct
NAME/KEY: CDS
LOCATION: (139)...(3406)
US-09-769-987-1

Alignment Scores:
Pred. No.: 8.55 Length: 6412
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.91% Indels: 0
DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-769-987-1 (1-6412)

OY 940 CysValHisArgAspLeuAlaIArgAsn 949
|||||
Db 2578 TGTCACCGTGTGCTGCTGCTGCAAC 2607

RESULT 90
US-09-982-610-17/C
Sequence 17, Application US/09982610
Patent No. US20020146420A1
GENERAL INFORMATION:
APPLICANT: Genentech, Inc.
Bennett, Brian D.
Goeddel, David
Lee, James M.
Matthews, William
Tsai, Siao Ping
Wood, William I.
TITLE OF INVENTION: PROTEIN TYROSINE KINASE AGONIST ANTIBODIES
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Winpatin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/982,610
FILING DATE: 17-Oct-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/446,648
FILING DATE: 1996-MAY-23
APPLICATION NUMBER: 08/222616
FILING DATE: 04-APR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M.
REGISTRATION NUMBER: 40,378
REFERENCE/DOCKET NUMBER: P0821P3PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1994
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 6827 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Single
TOPOLOGY: Linear
SEQUENCE DESCRIPTION: SEQ ID NO: 17:
US-09-982-610-17

Alignment Scores:
Pred. No.: 9.04 Length: 6827
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.91% Indels: 0
DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-982-610-17 (1-6827)

OY 942 HisArgAspLeuAlaIArgAsnIleu 951
|||||
Db 2671 CACAGAGCGTGTGCTGCGAGACATTCTG 2642

RESULT 91
US-09-982-610-45
Sequence 45, Application US/09982610
Patent No. US20020146420A1
GENERAL INFORMATION:
APPLICANT: Genentech, Inc.
Bennett, Brian D.
Goeddel, David
Lee, James M.
Matthews, William
Tsai, Siao Ping
Wood, William I.
TITLE OF INVENTION: PROTEIN TYROSINE KINASE AGONIST ANTIBODIES
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Winpatin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/982,610
FILING DATE: 17-Oct-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/446,648
FILING DATE: 1996-MAY-23
APPLICATION NUMBER: 08/222616
FILING DATE: 04-APR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M.
REGISTRATION NUMBER: 40,378
REFERENCE/DOCKET NUMBER: P0821P3PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1994
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 9108 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Single
TOPOLOGY: Linear
SEQUENCE DESCRIPTION: SEQ ID NO: 45:
US-09-982-610-45

Alignment Scores:
Pred. No.: 11.7 Length: 9108
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.91% Indels: 0
DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-982-610-45 (1-9108)

OY 942 HisArcAspLeuAlaIarGAsnIleLeu 951

DB 4065 CACAGGACCTGTGCTCGACATTTCTG 4094

RESULT 92

US-09-960-352-10875/C

; Sequence 10875, Application US/09960352

; Patent No. US20020137139A1

; GENERAL INFORMATION:

; APPLICANT: Warren, Wesley C.

; APPLICANT: Tao, Nengbing

; APPLICANT: Byatt, John C.

; APPLICANT: Mathialagan, Nagappan

; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND

; FILE REFERENCE: 16511.006/37-21(10298)C

; CURRENT APPLICATION NUMBER: US/09/960,352

; CURRENT FILING DATE: 2001-09-24

; NUMBER OF SEQ ID NOS: 15112

; SEQ ID NO. 10875

; LENGTH: 332

; TYPE: DNA

; ORGANISM: Bos taurus

; OTHER INFORMATION: Clone ID: 47-LIB188-004-Q1-E1-D4

US-09-960-352-10875

Alignment Scores:

Pred. No.: 6.35

Score: 9.00 Length: 332

Percent Similarity: 100.00% Matches: 9

Best Local Similarity: 100.00% Conservative: 0

Query Match: 0.82% Mismatches: 0

DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-960-352-10875 (1-332)

OY 1002 GlnSerAspValTrpSerPheGlyVal 1010

DB 259 CACAGCGACGTGTGCTTTGGGGGTG 233

RESULT 93

US-09-960-352-1177/C

; Sequence 1177, Application US/09960352

; Patent No. US20020137139A1

; GENERAL INFORMATION:

; APPLICANT: Warren, Wesley C.

; APPLICANT: Tao, Nengbing

; APPLICANT: Byatt, John C.

; APPLICANT: Mathialagan, Nagappan

; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND

; FILE REFERENCE: 16511.006/37-21(10298)C

; CURRENT APPLICATION NUMBER: US/09/960,352

; CURRENT FILING DATE: 2001-09-24

; NUMBER OF SEQ ID NOS: 15112

; SEQ ID NO. 1177

; LENGTH: 355

; TYPE: DNA

; ORGANISM: Bos taurus

; OTHER INFORMATION: Clone ID: 06-LIB188-012-Q1-E1-B9

US-09-960-352-1177

Alignment Scores:

Pred. No.: 6.74

Score: 9.00 Length: 355

Percent Similarity: 100.00% Matches: 9

Best Local Similarity: 100.00% Conservative: 0

Query Match: 0.82% Mismatches: 0

DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-960-352-1177 (1-355)

OY 1002 GlnSerAspValTrpSerPheGlyVal 1010

DB 259 CACAGCGACGTGTGCTTTGGGGGTG 233

RESULT 94

US-10-101-464A-332

; Sequence 332, Application US/10101464A

; Publication No. US20030046728A1

; GENERAL INFORMATION:

; APPLICANT: Strabala, Timothy

; APPLICANT: Nieuwenhuizen, Nicolaas

; APPLICANT: Higgins, Colleen M.

; TITLE OF INVENTION: Compositions Isolated from Plant Cells

; FILE REFERENCE: 11000.1020C2 and Their Use in the Modification of Plant Cell Signaling

; CURRENT APPLICATION NUMBER: US/10/101,464A

; CURRENT FILING DATE: 2002-03-18

; PRIOR APPLICATION NUMBER: 09/704,302

; PRIOR FILING DATE: 2000-11-01

; PRIOR APPLICATION NUMBER: 09/228,986

; PRIOR FILING DATE: 1999-01-12

; PRIOR APPLICATION NUMBER: 60/162,866

; PRIOR FILING DATE: 1999-11-01

; PRIOR APPLICATION NUMBER: PCT/US00/00724

; PRIOR FILING DATE: 2000-01-11

; NUMBER OF SEQ ID NOS: 989

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO. 332

; LENGTH: 363

; TYPE: DNA

; ORGANISM: Pinus radiata

US-10-101-464A-332

Alignment Scores:

Pred. No.: 6.88

Score: 9.00 Length: 363

Percent Similarity: 100.00% Matches: 9

Best Local Similarity: 100.00% Conservative: 0

Query Match: 0.82% Mismatches: 0

DB: 9 Gaps: 0

US-09-397-967-16 (1-1099) x US-10-101-464A-332 (1-363)

OY 961 AlaAspPheGlyLeuAlaIarGAsnIleLeu 969

DB 25 GCTGATTTGGGCTTGCCAGCGTTTG 51

RESULT 95

US-09-960-352-10878

; Sequence 10878, Application US/09960352

; Patent No. US20020137139A1

; GENERAL INFORMATION:

; APPLICANT: Warren, Wesley C.

; APPLICANT: Tao, Nengbing

; APPLICANT: Byatt, John C.

; APPLICANT: Mathialagan, Nagappan

; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND

; FILE REFERENCE: 16511.006/37-21(10298)C

; CURRENT APPLICATION NUMBER: US/09/960,352

; CURRENT FILING DATE: 2001-09-24

; NUMBER OF SEQ ID NOS: 15112

; SEQ ID NO. 10878

; LENGTH: 366

; TYPE: DNA

; ORGANISM: Bos taurus

; OTHER INFORMATION: Clone ID: 47-LIB188-007-Q1-E1-D4

US-09-960-352-10878

Alignment Scores:

Pred. No.: 6.93

Score: 9.00 Length: 366

Percent Similarity: 100.00% Matches: 9

Best Local Similarity: 100.00% Conservative: 0

Query Match: 0.82% Mismatches: 0

DB: 10 Gaps: 0

Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.82% Indels: 0
DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-960-352-10878 (1-366)

OY 1002 GlnSerAspValTrpSerPheGlyVal 1010
|||||
DB 289 CAGAGCGACGCTGTGCTTTGGGGTG 315

RESULT 96

US-09-960-352-2603
; Sequence 2603, Application US/09960352
; Patent No. US20020137139A1

GENERAL INFORMATION:

APPLICANT: Warren, Wesley C.

APPLICANT: Tao, Nengding

APPLICANT: Byatt, John C.

APPLICANT: Mathialagan, Nagappan

TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND

TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION

FILE REFERENCE: 16511.006/37-21(10298)C

CURRENT APPLICATION NUMBER: US/09/960,352

CURRENT FILING DATE: 2001-09-24

NUMBER OF SEQ ID NOS: 15112

SEQ ID NO 2603

LENGTH: 391

TYPE: DNA

ORGANISM: Bos taurus

OTHER INFORMATION: Clone ID: 12-LIB188-021-Q1-E1-C7

US-09-960-352-2603

Alignment Scores:

Pred. No.: 7.35 Length: 391
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.82% Indels: 0
DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-960-352-2603 (1-391)

OY 1002 GlnSerAspValTrpSerPheGlyVal 1010
|||||
DB 291 CAGAGCGACGCTGTGCTTTGGGGTG 317

RESULT 97

US-09-960-352-4293

; Sequence 4293, Application US/09960352

; Patent No. US20020137139A1

GENERAL INFORMATION:

APPLICANT: Warren, Wesley C.

APPLICANT: Tao, Nengding

APPLICANT: Byatt, John C.

APPLICANT: Mathialagan, Nagappan

TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND

TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION

FILE REFERENCE: 16511.006/37-21(10298)C

CURRENT APPLICATION NUMBER: US/09/960,352

CURRENT FILING DATE: 2001-09-24

NUMBER OF SEQ ID NOS: 15112

SEQ ID NO 4293

LENGTH: 392

TYPE: DNA

ORGANISM: Bos taurus

OTHER INFORMATION: Clone ID: 19-LIB188-015-Q1-E1-E3

US-09-960-352-4293

Alignment Scores:

Pred. No.: 7.37 Length: 392
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0

Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.82% Indels: 0
DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-960-352-4293 (1-392)

OY 1002 GlnSerAspValTrpSerPheGlyVal 1010
|||||
DB 291 CAGAGCGACGCTGTGCTTTGGGGTG 317

RESULT 98

US-09-960-352-13765
; Sequence 13765, Application US/09960352
; Patent No. US20020137139A1

GENERAL INFORMATION:

APPLICANT: Warren, Wesley C.

APPLICANT: Tao, Nengding

APPLICANT: Byatt, John C.

APPLICANT: Mathialagan, Nagappan

TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND

TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION

FILE REFERENCE: 16511.006/37-21(10298)C

CURRENT APPLICATION NUMBER: US/09/960,352

CURRENT FILING DATE: 2001-09-24

NUMBER OF SEQ ID NOS: 15112

SEQ ID NO 13765

LENGTH: 401

TYPE: DNA

ORGANISM: Bos taurus

OTHER INFORMATION: Clone ID: 59-LIB2809-008-Q1-E1-G12

US-09-960-352-13765

Alignment Scores:

Pred. No.: 7.52 Length: 401
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.82% Indels: 0
DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-960-352-13765 (1-401)

OY 660 ProPheIleLysLeuSerAspProGly 668
|||||
DB 243 CCTTCATCAAGCTCAGTGCCTGCGC 269

RESULT 99

US-09-960-352-14442

; Sequence 14442, Application US/09960352

; Patent No. US20020137139A1

GENERAL INFORMATION:

APPLICANT: Warren, Wesley C.

APPLICANT: Tao, Nengding

APPLICANT: Byatt, John C.

APPLICANT: Mathialagan, Nagappan

TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND

TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION

FILE REFERENCE: 16511.006/37-21(10298)C

CURRENT APPLICATION NUMBER: US/09/960,352

CURRENT FILING DATE: 2001-09-24

NUMBER OF SEQ ID NOS: 15112

SEQ ID NO 14442

LENGTH: 402

TYPE: DNA

ORGANISM: Bos taurus

OTHER INFORMATION: Clone ID: 62-LIB188-002-Q1-E1-H10

US-09-960-352-14442

Alignment Scores:

Pred. No.: 7.54 Length: 402
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.82% Indels: 0
 DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-960-352-14442 (1-402)

OY 1002 GlnSerAspValTrpSerPheGlyVal 1010

DB 301 CAGACGACGCTGTCTTTGGGGTG 327

RESULT 100

US-09-918-995-16302

Sequence 16302, Application US/09918995

Publication No. US20030073623A1

GENERAL INFORMATION:

APPLICANT: Hyseq, Inc.

TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED

TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES

FILE REFERENCE: 20411-756

CURRENT APPLICATION NUMBER: US/09/918,995

PRIOR FILING DATE: 2001-07-30

PRIOR APPLICATION NUMBER: US/09/235,076

NUMBER OF SEQ ID NOS: 38054

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 16302

LENGTH: 415

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc_feature

LOCATION: (1)...(415)

OTHER INFORMATION: n = A,T,C or G

Alignment Scores:

Pred. No.: 7.75

Score: 9.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 0.82%

DB: 9

US-09-397-967-16 (1-1099) x US-09-918-995-16302 (1-415)

OY 941 ValHisArgAspLeuAlaAlaArgAsn 949

DB 133 GTACACAGGAGCTGGCGCTCGGAGC 159

Search completed: April 28, 2003, 23:26:17
 Job time : 451 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: April 28, 2003, 18:58:45 ; Search time 137 seconds
(without alignments)
2460.130 Million cell updates/sec

Title: US-09-397-967-16

Perfect score: 1099

Sequence: 1 MAPPERPPLIPORSCSLSS.....RPAPATLSPOLDPLNMRGPG 1099

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 153338381 residues

Word size: 1

Total number of hits satisfying chosen parameters: 878600

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 150 summaries

Command line parameters:

-MODE=frame_p2n.model -DEV=xlp
-O=/cgn2.1/USPTO.spool/US09397967/runat.26042003.182927.11236/app_query.fasta.1.1287
-DB=Issued Patents.NA -QFMT=fastexp -SUPER=oligo.rml -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=150 -DOCCALIGN=200 -THR.SCORE=quality -THR.MIN=1 -ALIGN=100 -MODE=LOCAL
-OUTFMT=PCO -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09397967.ecn1.1.49.0runat.26042003.182927.11236 -NCPU=6 -ICPU=3
-NO_XLPEXT -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database:

Issued Patents.NA:*
1: /cgn2.6/plodata/2/lna/5A.COMB.seq:*
2: /cgn2.6/plodata/2/lna/5B.COMB.seq:*
3: /cgn2.6/plodata/2/lna/6A.COMB.seq:*
4: /cgn2.6/plodata/2/lna/6B.COMB.seq:*
5: /cgn2.6/plodata/2/lna/PCRTS.COMB.seq:*
6: /cgn2.6/plodata/2/lna/Backfillseq1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	180	16.4	4016	5	PCRT-US95-08354A-1
2	62	5.6	3807	1	US-08-357-598-1
3	62	5.6	3807	2	US-09-003-289-1
4	62	5.6	3807	5	PCRT-US95-16435-1
5	16	1.5	3435	4	US-09-046-158A-21
6	16	1.5	3435	4	US-08-446-038B-2
7	16	1.5	3435	1	US-08-446-010B-2
8	16	1.5	3435	1	US-08-805-445-2
9	16	1.5	3435	2	US-08-064-067D-2
10	16	1.5	3435	2	US-09-066-208-2
11	16	1.5	3435	4	US-08-980-080-3
12	16	1.5	3629	1	US-08-097-997A-8

13	16	1.5	3629	3	US-08-665-574C-8	Sequence 8, Appl1
14	16	1.5	3629	4	US-08-946-994-8	Sequence 8, Appl1
15	16	1.5	4482	2	US-08-567-508C-1	Sequence 1, Appl1
16	16	1.5	4482	3	US-09-196-480-1	Sequence 1, Appl1
17	13	1.2	41	1	US-08-481-003-9	Sequence 9, Appl1
18	13	1.2	41	3	US-08-485-598-9	Sequence 9, Appl1
19	13	1.2	45	1	US-08-481-003-8	Sequence 8, Appl1
20	13	1.2	45	3	US-08-485-598-8	Sequence 8, Appl1
21	12	1.1	151	1	US-08-222-616-11	Sequence 11, Appl1
22	12	1.1	151	5	PCRT-US95-04228-11	Sequence 11, Appl1
23	12	1.1	151	4	US-08-446-648-11	Sequence 11, Appl1
24	12	1.1	738	2	US-08-604-989A-8	Sequence 8, Appl1
25	12	1.1	1398	2	US-08-604-989A-9	Sequence 9, Appl1
26	12	1.1	1521	2	US-08-604-989A-10	Sequence 10, Appl1
27	12	1.1	1713	4	US-09-741-154-1	Sequence 1, Appl1
28	12	1.1	1442	4	US-08-604-989A-11	Sequence 11, Appl1
29	12	1.1	1987	2	US-08-876-882-1	Sequence 1, Appl1
30	12	1.1	1987	4	US-09-315-928-1	Sequence 1, Appl1
31	12	1.1	2000	4	US-08-426-509A-1	Sequence 1, Appl1
32	12	1.1	2000	5	PCRT-US95-05008-1	Sequence 1, Appl1
33	12	1.1	2820	1	US-08-162-809-5	Sequence 1, Appl1
34	12	1.1	2862	2	US-08-449-645A-10	Sequence 10, Appl1
35	12	1.1	2862	2	US-08-702-367A-10	Sequence 10, Appl1
36	12	1.1	2962	5	PCRT-US95-04681-10	Sequence 10, Appl1
37	12	1.1	2982	1	US-08-348-143-2	Sequence 2, Appl1
38	12	1.1	2982	4	US-08-571-785-2	Sequence 2, Appl1
39	12	1.1	2982	4	US-09-192-435-2	Sequence 2, Appl1
40	12	1.1	2982	4	US-09-558-340-2	Sequence 2, Appl1
41	12	1.1	3105	4	US-08-542-635-1	Sequence 1, Appl1
42	12	1.1	3116	2	US-08-449-645A-14	Sequence 14, Appl1
43	12	1.1	3116	2	US-08-702-367A-14	Sequence 14, Appl1
44	12	1.1	3116	5	PCRT-US95-04681-14	Sequence 14, Appl1
45	12	1.1	3133	1	US-08-162-809-1	Sequence 1, Appl1
46	12	1.1	3150	4	US-09-166-350-23	Sequence 23, Appl1
47	12	1.1	3348	4	US-08-222-616-34	Sequence 34, Appl1
48	12	1.1	3348	4	US-08-446-648-34	Sequence 34, Appl1
49	12	1.1	3348	5	PCRT-US95-04228-34	Sequence 34, Appl1
50	12	1.1	3546	1	US-08-162-809-9	Sequence 9, Appl1
51	12	1.1	3591	1	US-08-162-809-13	Sequence 13, Appl1
52	12	1.1	3792	2	US-08-466-537A-100	Sequence 100, App
53	12	1.1	3766	1	US-08-162-809-7	Sequence 7, Appl1
54	12	1.1	3969	1	US-08-436-004-5	Sequence 5, Appl1
55	12	1.1	3969	2	US-08-222-616-23	Sequence 23, Appl1
56	12	1.1	3969	2	US-08-436-054-5	Sequence 5, Appl1
57	12	1.1	3969	4	US-08-446-648-23	Sequence 23, Appl1
58	12	1.1	3969	5	PCRT-US95-04228-23	Sequence 23, Appl1
59	12	1.1	3969	5	PCRT-US95-08812-5	Sequence 5, Appl1
60	12	1.1	4027	1	US-08-348-143-3	Sequence 3, Appl1
61	12	1.1	4027	1	US-08-348-143-4	Sequence 4, Appl1
62	12	1.1	4027	1	US-08-571-785-3	Sequence 3, Appl1
63	12	1.1	4027	1	US-08-571-785-4	Sequence 4, Appl1
64	12	1.1	4027	4	US-09-192-435-3	Sequence 3, Appl1
65	12	1.1	4027	4	US-09-192-435-4	Sequence 4, Appl1
66	12	1.1	4027	4	US-09-558-340-3	Sequence 3, Appl1
67	12	1.1	4027	4	US-09-558-340-4	Sequence 4, Appl1
68	12	1.1	4049	1	US-08-162-809-17	Sequence 17, Appl1
69	12	1.1	4097	4	US-08-162-809-11	Sequence 11, Appl1
70	12	1.1	4304	4	US-08-368-776A-1	Sequence 1, Appl1
71	12	1.1	4304	5	PCRT-US96-00419-1	Sequence 1, Appl1
72	12	1.1	4508	5	PCRT-US93-06251-34	Sequence 34, Appl1
73	12	1.1	4529	2	US-08-449-645A-16	Sequence 16, Appl1
74	12	1.1	4529	5	US-08-702-367A-16	Sequence 16, Appl1
75	12	1.1	4529	2	PCRT-US95-04681-16	Sequence 16, Appl1
76	12	1.1	5816	4	US-08-857-076-11	Sequence 11, Appl1
77	12	1.1	16389	4	US-09-741-154-3	Sequence 3, Appl1
78	12	1.1	147	2	US-08-222-616-7	Sequence 7, Appl1
79	12	1.1	147	2	US-08-876-882-7	Sequence 7, Appl1
80	12	1.1	147	4	US-08-446-648-7	Sequence 7, Appl1
81	12	1.1	147	5	PCRT-US95-04228-7	Sequence 7, Appl1
82	12	1.1	1590	1	US-08-278-089A-3	Sequence 3, Appl1
83	12	1.1	1590	1	US-07-934-393B-3	Sequence 3, Appl1
84	12	1.1	1601	1	US-08-838-997A-3	Sequence 3, Appl1
85	12	1.1	1601	2	US-08-838-997A-3	Sequence 3, Appl1

86	11	1.0	3056	1	US-08-162-809-21
87	11	1.0	3059	1	US-08-162-809-3
88	11	1.0	3125	1	US-08-162-809-19
89	11	1.0	3132	3	US-08-167-919A-9
90	11	1.0	3132	3	US-08-115-106-9
91	11	1.0	3162	2	US-08-449-645A-12
92	11	1.0	3162	2	US-08-702-367A-12
93	11	1.0	3162	5	PCT-US95-04681-12
94	11	1.0	3254	1	US-08-162-809-15
95	11	1.0	3306	2	US-08-469-537A-102
96	11	1.0	4138	1	US-08-323-474-1
97	11	1.0	4138	5	PCT-US93-06099-1
98	11	1.0	4165	1	US-08-442-248-1
99	11	1.0	4165	1	US-08-440-815-1
100	11	1.0	4165	4	US-08-486-449-1
101	11	1.0	4175	1	US-07-934-393B-1
102	11	1.0	4175	1	US-08-778-089A-1
103	11	1.0	4175	2	US-08-638-957A-1
104	11	1.0	4176	1	US-08-778-089A-5
105	11	1.0	4176	2	US-08-638-957A-5
106	11	1.0	4322	1	US-08-673-789-1
107	11	0.9	284	5	PCT-US92-02750-5
108	10	0.9	510	1	US-08-810-116-5
109	10	0.9	510	2	US-07-930-548A-5
110	10	0.9	1894	4	US-07-912-122-3
111	10	0.9	1894	5	PCT-US99-06404-3
112	10	0.9	2301	1	US-08-306-651B-23
113	10	0.9	2301	5	PCT-US93-06251-78
114	10	0.9	2820	1	US-08-286-305A-4
115	10	0.9	2820	2	US-08-441-104A-4
116	10	0.9	2820	4	US-09-417-881A-4
117	10	0.9	3120	4	US-08-222-517-22
118	10	0.9	3120	4	US-08-446-448-22
119	10	0.9	3120	5	PCT-US95-04228-22
120	10	0.9	3120	5	PCT-US93-08133-993-1
121	10	0.9	3453	1	US-07-917-451-1
122	10	0.9	3453	1	US-07-946-507-1
123	10	0.9	3453	1	US-08-252-517-1
124	10	0.9	3453	1	US-08-262-397A-1
125	10	0.9	3453	1	US-08-601-891-1
126	10	0.9	3453	2	US-09-021-324-1
127	10	0.9	3453	5	PCT-US92-02750-1
128	10	0.9	3453	5	PCT-US92-05893-1
129	10	0.9	3453	5	PCT-US99-09480-1
130	10	0.9	3475	1	US-08-222-299-3
131	10	0.9	3475	2	US-08-424-878-3
132	10	0.9	3475	5	PCT-US99-03718-3
133	10	0.9	3476	5	US-08-183-211-1
134	10	0.9	3476	5	PCT-US99-00176A-1
135	10	0.9	3501	1	US-07-917-451-3
136	10	0.9	3501	1	US-08-252-517-3
137	10	0.9	3501	1	US-07-906-397A-3
138	10	0.9	3501	1	US-08-601-891-3
139	10	0.9	3501	2	US-09-021-324-3
140	10	0.9	3501	5	PCT-US99-05893-3
141	10	0.9	3501	5	PCT-US99-09480-1
142	10	0.9	3521	1	US-08-222-299-1
143	10	0.9	3521	2	US-08-424-878-1
144	10	0.9	3521	5	PCT-US95-03718-1
145	10	0.9	3561	1	US-08-097-997A-12
146	10	0.9	3561	3	US-08-665-574C-12
147	10	0.9	3561	4	US-08-466-994-12
148	10	0.9	4054	1	US-08-160-195-15
149	10	0.9	4054	1	US-08-477-329-35
150	10	0.9	4054	1	US-08-477-329-35

ALIGNMENTS

[illegible]

```

RESULT 1
PCT-US95-08354A-1
; Sequence 1, Application PC/TUS9508354A
; GENERAL INFORMATION:

```

```

APPLICANT: Temple University - Of The
APPLICANT: Commonwealth System of Higher Education
TITLE OF INVENTION: KINASE PROTEIN TYROSINE
TITLE OF INVENTION: KINASE AND DNA ENCODING THE SAME
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seidel, Conda, Lavorgna
ADDRESSEE: 6 Monaco, P.C.
STREET: Suite 1800, Two Penn Center
CITY: Philadelphia
STATE: Pennsylvania
COUNTY: U.S.A.
ZIP: 19102
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT-US95/08354A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/272,368
FILING DATE: 8 July 1994
ATTORNEY/AGENT INFORMATION:
NAME: Monaco, Daniel A.
REGISTRATION NUMBER: 30,480
REFERENCE/DOCKET NUMBER: 6056-203 PC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-8383
TELEFAX: (215) 568-5549
INFORMATION FOR SEQ ID NO.: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4016 nucleotides
TYPE: nucleic acid
STRANDEDNESS: single stranded
TOPOLOGY: linear
PCT-US95-08354A-1

Alignment Scores:
Pred. No.:      8,44e-154      Length:      4016
Score:          180.00         Matches:     529
Percent Similarity: 97.24%    Conservative: 0
Best local Similarity: 97.24% Mismatches:   9
Query Match:     16.38%       Indels:     15
Db:              5           Gaps:       0

US-09-397-967-16 (1-1099) x PCT-US95-08354A-1 (1-4016)

Qy      535  ArGaTgGluVaValAspGlyGluTrnHisAspSergLuuValLeuLeuLySaValMetasp      554
        |||||||
Db      2257 AGGGGGAGGCGTGGAATGGTGAGACACATGACTCGGAAGTCCTCCTTAAGGTCAATGGAC      2316

Qy      555  SerArghisArgasnCySwetGtUserPhleUgluaAlaSerLeuUelSergLuvAl      574
        |||||||
Db      2317 TCCAGACATCGGAACCTGCATGGAAGCTTTTCTGGAAGCCGAAAGCTTATGAGCCAASTA      2376

Qy      575  SerIyrProHsisleuValileuHhisGlyValIcysMetaIagIyaSPserIlemetVal      594
        |||||||
Db      2377 TCCTACCAGCACCTGGGTGTTACTGACACAGCGCTGTGCATGGCTGGAGACAGCATATCGTG      2436

Qy      595  GlnglUpheValIyrlengIyalalaleasPmetTyrlLeuargLySarGlyHhisLeuVal      614
        |||||||
Db      2437 CAGCAATTGTGTATCAGAGAACAATTGACATGTACTCGCAACACGTGGCCACCTGGTG      2496

Qy      615  SerAlaserTrpylsleuGlInvalTrnIlysGlnLeuAlaTyrrAlaLeuasnTyrlengLu      634
        |||||||
Db      2497 TCAGCACACCTGGAAACGCAAGGTGACCACACACACTGGCATATGCCCCTTTAACTACTTGGAG      2556

Qy      635  AspLyScGlyLeuPProHsisGlyAsnValSerAlaIryLySaValleuLeuAlaIarGluGly      654
        |||||||
Db      2557 GACAAGGCCCTTCTCATCGGCAACGTCCTCATGACGGAAGAGTGCTCCTGCTGCTGACGGG      2616

```



```

QY 655 G1yAspG1yasnProPheIleIleuSerAspProG1yValSerProThrValLeu 674
DB 2617 GGtGATGGGAAATCCACCTTTCAATAGAGTACGTCGTCAGTCCACCTGTCG 2676
QY 675 SerLeuG1uMetLeuThrAspArgIleProTrpValAlaProG1uG1uG1uG1uAla 694
DB 2677 AGCCTGAAATGCTCAGACCCAGAAATACCTGGGTGGCCCGCAATGTCTCCAGGAGGCT 2736
QY 695 G1uThrLeuG1uG1uG1uAlaAspLysTrpG1yValAlaThrThrTrpG1uValPhe 714
DB 2737 CACACACTGCTGCTGGAGGTGACAAAGTGGGCTTTGGAGCCACAGTGGGAGGTGT- 2795
QY 715 G1uArg-G1yProAlaHisIleThrSerLeuG1uProAlaLysLysLeuLysPheTrpG1 734
DB 2796 CACGGGGGAGCCCGCCACATCCCTCGCGAGGCCCGCCAAAACCTGAAGTCTATGA 2855
QY 734 uAspG1uG1uG1uProAlaLeuLysTrpThrG1uLeuAlaG1yLeuIleThrG1uG1y 754
DB 2856 GGACCAAGGAGACAGCTCCGCTCTCAATGACAGAACTGGCGGAGCTTATCACACAGTG 2915
QY 754 sMetAlaTrpAspProG1yArgArgProSerPheArgAlaIleLeuArgAspLeuSngI 774
DB 2916 CATGGGCTATGATCTGCGCGCGCGCTCTCTCCAGCTATCTAGAGACTCAACGG 2975
QY 774 ILeuIleThrSerAspTrpG1uLeuLeuSerAspProThrProG1yIleProSerProAr 794
DB 2976 CCTCATTTACATCAGATTAGAGAGCTCTCCA-GACCCACACCTGGCATCCGAGTCTCTG 3034
QY 794 gAspG1uLeuG1yValAlaG1yAlaG1uLeuTrpAlaG1yG1uAspProAlaIlePheG1 814
DB 3035 AGATGAGCTGTGCG--T--GGCGCCCAAGCTATGCTCCAGGACCCCGCATATAC-GA 3089
QY 814 uG1uArgHisLeuLysTrpIleSerLeuG1uG1yG1yG1uAspPheG1SerValG1uLe 834
DB 3090 GGAGAGACACCTTAATATATCTTGTGTGGGCAAGGCACTTTGGACAGCTGAGCT 3149
QY 834 uGysArgTrpAspProLeuG1yAspAsnThrG1yProLeuValAlaValLysG1uLeuG1 854
DB 3150 GTGCCCTTATGACCCCC--TG-GACAAATGAGGAGCCCGCTGTCGAGTGAACACACTACA 3206
QY 854 nHisSerVal-PROARGING1uArgAspPheG1uArgG1uIleG1uIleLeuLysAlaI 874
DB 3207 GCACAGCCG-GCCACAGCCAGAGAGGACTTCCAGCGGAGATTGAGATCTTAAGGCTC 3265
QY 874 euHisSerAspPheIleValLysTrpArgG1yValSerTrpG1yProG1uArgG1uSerL 894
DB 3266 TCCACAGGACTTCAATCCGTCAGATACCGGGAGTCAAGTATGGCCAGAGTCCCAAGGCC 3325
QY 894 euAlaLeuValMetG1uTrpTrpLeuProSerG1yCysLeuArgAspLeu-LeuG1uArgHis 913
DB 3326 TCCGGTGTGTGATGAGTACCTGCCAGCGGCTGCTGGAGA-CTTCTCGACAGCGCAT 3384
QY 914 --ArgG1y-LeuHisThrAspArgLeuLeuPheAlaTrpG1uIleCysLysG1uMet 932
DB 3385 CGCGGGGGCCCTCAGACAGCCAGCGCTACTGCTTCCGTCGCAATCTGCAAGGCGATG 3444
QY 933 G1uTrpLeuG1yAlaArgArgCysValHisArgAspLeuAlaIleArgAsnIleLeuVal 952
DB 3445 GAGTACCTGGGTGCGCGGCTGCTGACACCTGACCTGGGCTGCCGCAACATCTTGGTG 3504
QY 953 G1uSerG1uAlaHisValLysIleAlaAspPheG1yLeuAlaLysLeuLeuProLeuG1y 972
DB 3505 GAGACGAGGCTCATGTATGAATCGCGGACTTCCGCTCCGTAACCTGCTGCCCTGGGA 3564
QY 973 LysAspTrpTrpValValArgG1uProG1yG1uSerProIlePheTrpTrpAlaProG1y 992
DB 3565 AAGAGCTACTAGTGTGCGGAGCTGCCAAGCCCATCTTTGGTATAGCCCGGAG 3624
QY 993 SerLeuSerAspAsnIlePheSerArgIleSerAspValTrpSerPheG1yValValLeu 1012
DB 3625 TCCCTATCTGACAACTTCTCCCGCAATCTGACGTGTGAGAGCTTGGAGTGTGTG 3684

```

```

QY 1013 TyrG1uLeuPheThrTrpTrpCysAspLysSerCysSerProSerAlaG1uPheLeuArgMet 1032
DB 3685 TACGAGCTCTTCACTACCTACGCGACAGAGAGTGTAGCCCATCCGCTGAGTCTCCGCGCATG 3744
QY 1033 MetG1yProG1uArgG1uG1yProProLeuG1yArgLeuLeuG1uLeuAlaG1uG1y 1052
DB 3745 ATGGGGCTGAGCGGTAAAGAGCCCGCGCTCTGCGGCTGTGAAGCTGTGGGAGAGGGC 3804
QY 1053 ArgArgLeuProProProProThrCysProThrG1uValG1uG1uMetG1uLeuG1y 1072
DB 3805 CGAGCGCTCCACACACTCCACCTCCACCTGCGCCAGAGAGTTCAGAGCTCATGAGCTGTGC 3864

RESULT 2
US-08-357-598-1
? Sequence 1, Application US/08357598
? Patent No. 5705625
? GENERAL INFORMATION:
? APPLICANT: CLYIN, CURT I.
? APPLICANT: Small, Donald
? TITLE OF INVENTION: NOVEL PROTEIN TYROSINE KINASE, JAK3
? NUMBER OF SEQUENCES: 12
? CORRESPONDENCE ADDRESS:
? ADDRESS: Fish & Richardson P.C.
? STREET: 4225 Executive Square, Suite 1400
? CITY: La Jolla
? STATE: CA
? COUNTRY: USA
? ZIP: 92037
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patent Release #1.0, Version #1.30
? CURRENT APPLICATION DATA: US/08/357, 598
? APPLICATION NUMBER: US/08/357, 598
? FILING DATE: 15-DEC-1994
? CLASSIFICATION: 435
? ATTORNEY/AGENT INFORMATION:
? NAME: Haile, Lisa A.
? REGISTRATION NUMBER: 38,347
? REFERENCE/DOCKET NUMBER: 07265/033001
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 619/678-5070
? TELEFAX: 619/678-5099
? INFORMATION FOR SEQ ID NO: 1:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 3807 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: DNA
? US-08-357-598-1

Alignment Scores:
Pred. No.: 1,24e-46 Length: 3807
Score: 62.00 Matches: 96
Percent Similarity: 97.96% Conserves: 0
Best Local Similarity: 97.96% Mismatches: 1
Query Match: 5.64% Indels: 2
DB: 1 Gaps: 0

US-09-397-967-16 (1-1099) x US-08-357-598-1 (1-3807)
QY 938 ArgArgCysValHisArgAspLeuAlaIleArgAsnIleLeuValG1uSerG1uAlaHis 957
DB 2965 CGCGGCTGCGTGCACCGGAGCTGGCGCCGGAACATCTCTGTGGAGAGCGAGCACAC 3044
QY 958 ValLysIleAlaAspPheG1yLeuAlaLysLeuLeuPro-LeuG1yLysAspTrpTrpVal 977
DB 3045 GTCAAGATGCTGACTTGGGCTGAGCTAGCTGTGCGGCTTGC-AAAAGCTACTACGT 3103
QY 977 lValArgG1uProG1yG1uSerProIlePheTrpTrpAlaProG1uSerLeuSerAspAs 997

```

Db 3104 GGTCCGGAGCCAGCCAGACCCCATTTTCTGTATGCCCCCGAATCCCTCTGGACAA 3163
OY 997 nllpHeserArGlnserAspValTrrpSerPheGlyValValleuTyrGluLeuPheTh 1017
Db 3164 CATTTCTCTCGCCAGTCAGACGCTGTGAGGTGGGGTGTCTCTGTACAGACTCTTAC 3223
OY 1017 rTyrCyAspLysSerCysSerProSerAlaGluPheLeuArgMetGly 1034
Db 3224 CTACTGGACAAAGCTGACAGCCCTCGCGCGAGTCTCTCGGATGATGGGA 3275

RESULT 3

US-09-003-289-1
; Sequence 1, Application US/09003289
; Patent No. 5916792
; GENERAL INFORMATION:
; APPLICANT: CIVIL, Curt I.
; APPLICANT: Small, Donald
; TITLE OF INVENTION: NOVEL PROTEIN TYROSINE KINASE, JAK3
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/003,289
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/357,598
; FILING DATE: 15-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Hallie, Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07265/033001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5099
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3807 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-09-003-289-1

Alignment Scores:
Pred. No.: 1.24e-46 Length: 3807
Score: 62.00 Matches: 96
Percent Similarity: 97.96% Conservative: 0
Best Local Similarity: 97.96% Mismatches: 1
Query Match: 5.64% Indels: 2
Gaps: 0

US-09-397-967-16 (1-1099) x US-09-003-289-1 (1-3807)

OY 938 ArgArGysValHisArGAspLeuAlaAlaArgAsnIleLeuValGluSerGluAlaHis 957
Db 2985 CGCCCTGCGTGCACCGCAGCTGGCCCGCAACATCTCTGTGAGAGCGAGGACAC 3044
OY 958 ValLysIleAlaAspPheGlyLeuAlaLysLeuLeuPro-LeuGlyLysAspTyrTyrVa 977
Db 3045 GTCAGAGTCGCTGACTGCGCTAGCTAGCTGCTGCCCTTGAC-AAAGACTACTACGT 3103
OY 977 lValAlrGgluPrGlnserProIlePheTrrPyrAlaPrGluSerLeuSerAspAs 997

Db 3104 GGTCCGGAGCCAGCCAGACCCCATTTTCTGTATGCCCCCGAATCCCTCTGGACAA 3163
OY 997 nllpHeserArGlnserAspValTrrpSerPheGlyValValleuTyrGluLeuPheTh 1017
Db 3164 CATTTCTCTCGCCAGTCAGACGCTGTGAGGTGGGGTGTCTCTGTACAGACTCTTAC 3223
OY 1017 rTyrCyAspLysSerCysSerProSerAlaGluPheLeuArgMetGly 1034
Db 3224 CTACTGGACAAAGCTGACAGCCCTCGCGCGAGTCTCTCGGATGATGGGA 3275

RESULT 4

PCT-US95-16435-1
; Sequence 1, Application PC/TUS9516435
; GENERAL INFORMATION:
; APPLICANT: The Johns Hopkins University School of Medicine
; TITLE OF INVENTION: NOVEL PROTEIN TYROSINE KINASE, JAK3
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/16435
; FILING DATE: 15-DEC-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Hallie, Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07265/033W01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5099
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3807 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
PCT-US95-16435-1

Alignment Scores:
Pred. No.: 1.24e-46 Length: 3807
Score: 62.00 Matches: 96
Percent Similarity: 97.96% Conservative: 0
Best Local Similarity: 97.96% Mismatches: 1
Query Match: 5.64% Indels: 2
Gaps: 0

US-09-397-967-16 (1-1099) x PCT-US95-16435-1 (1-3807)

OY 938 ArgArGysValHisArGAspLeuAlaAlaArgAsnIleLeuValGluSerGluAlaHis 957
Db 2985 CGCCCTGCGTGCACCGCAGCTGGCCCGCAACATCTCTGTGAGAGCGAGGACAC 3044
OY 958 ValLysIleAlaAspPheGlyLeuAlaLysLeuLeuPro-LeuGlyLysAspTyrTyrVa 977
Db 3045 GTCAGAGTCGCTGACTGCGCTAGCTAGCTGCTGCCCTTGAC-AAAGACTACTACGT 3103
OY 977 lValAlrGgluPrGlnserProIlePheTrrPyrAlaPrGluSerLeuSerAspAs 997
Db 3104 GGTCCGGAGCCAGCCAGACCCCATTTTCTGTATGCCCCCGAATCCCTCTGGACAA 3163
OY 997 nllpHeserArGlnserAspValTrrpSerPheGlyValValleuTyrGluLeuPheTh 1017

Db 3164 CATCTCTCTGCGAGCTGAGAGCTGTGGAGCTTGGGGTCTCTGTAACAGCTCTTAC 3223
OY 1017 rTyrCysAspIysSerCysSerProSerAlaGlnPheLeuArgMetHetGly 1034
Db 3224 CTACTGCACAAAGCTGACGCCCTCCCTGGCGAGTCTCTGGATGATGGGA 3275

RESULT 5
US-09-046-158A-21

; Sequence 21, Application US/09046158A
; Patent No. 6187552
; GENERAL INFORMATION:
; APPLICANT: Roberda, Steven L.
; TITLE OF INVENTION: METHOD FOR IDENTIFYING INHIBITORS OF
; TITLE OF INVENTION: JAK2/CYTOKINE RECEPTOR BINDING
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pharmacia & Upjohn Co., Intellectual Property
; STREET: 301 Henrietta Street
; CITY: Kalamazoo
; STATE: MI
; COUNTRY: USA
; ZIP: 49001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/046,158A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Darnley Jr., James D.
; REGISTRATION NUMBER: 33,673
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 616/833-2210
; TELEFAX: 616/833-8897
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3435 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-09-046-158A-21

Alignment Scores:

Pred. No.:	6.9e-05	Length:	3435
Score:	16.00	Matches:	16
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.46%	Indels:	0
DB:	4	Gaps:	0

US-09-397-967-16 (1-1099) x US-09-046-158A-21 (1-3435)

OY 1003 SerAspValITrpSerPheGlyValIleuTyrGluLeuPheThrTyr 1018
Db 3106 TCAGATGTTGGAGCTTGGAGTGTCTGTATGAACCTTTTCACATAC 3153

RESULT 6
US-08-446-038B-2

; Sequence 2, Application US/08446038B
; Patent No. 5658791
; GENERAL INFORMATION:
; APPLICANT: Wilks, Andrew F.; Ziemlecki, Andrew;
; APPLICANT: Harpur, Alisa
; TITLE OF INVENTION: No. 5658791el Protein Tyrosine Kinase
; NUMBER OF SEQUENCES: 23

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felte & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022

COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,038B
; FILING DATE: 19-MAY-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/064,067
; FILING DATE: 30-Jun-1993
; APPLICATION NUMBER: PCT/US91/08889
; FILING DATE: 26-No. 5658791-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: Australian PK3594/90
; FILING DATE: 28-No. 5658791-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: Australian 88229/91
; FILING DATE: 27-No. 5658791-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5658791man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5244
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-688-9200
; TELEFAX: 212-838-3884
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3495 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: nucleic acid
US-08-446-038B-2

Alignment Scores:

Pred. No.:	7e-05	Length:	3495
Score:	16.00	Matches:	16
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.46%	Indels:	0
DB:	1	Gaps:	0

US-09-397-967-16 (1-1099) x US-08-446-038B-2 (1-3495)

OY 1003 SerAspValITrpSerPheGlyValIleuTyrGluLeuPheThrTyr 1018
Db 2695 TCAGATGTGTGAGAGCTTGGAGTGTCTATACGAACCTTTTCACATAC 2742

RESULT 7
US-08-446-010B-2

; Sequence 2, Application US/08446010B
; Patent No. 571618
; GENERAL INFORMATION:
; APPLICANT: Wilks, Andrew F.; Ziemlecki, Andrew;
; APPLICANT: Harpur, Alisa
; TITLE OF INVENTION: No. 571618el Protein Tyrosine Kinase
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felte & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/446,010B
FILING DATE: 19-May-1995
CLASSIFICATION: 433
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/446,038
FILING DATE: 19-May-1995
PRIOR APPLICATION DATA: 08/064,067
APPLICATION NUMBER: 30-Jun-1993
FILING DATE: 26-Jun-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US91/08889
FILING DATE: 26-Jun-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: Australian PK3594/90
FILING DATE: 28-Jun-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: Australian 88229/91
FILING DATE: 27-Jun-1991
APPLICATION NUMBER: Australian 88229/91
FILING DATE: 27-Jun-1991
ATTORNEY/AGENT INFORMATION:
NAME: Baer, Madeline F.
REGISTRATION NUMBER: 36,437
REFERENCE/DOCKET NUMBER: LUD 5244.3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-688-9200
TELEFAX: 212-838-3884
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 3495 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: nucleic acid
US-08-446-010B-2

Alignment Scores:
Pred. No.: 7e-05 Length: 3495
Score: 16.00 Matches: 16
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.46% Indels: 0
Gaps: 0
DB: 1

US-09-397-967-16 (1-1099) x US-08-446-010B-2 (1-3495)

QY 1003 SeraspvAITPserPheglYvalLeuTYrgluLeuPheThrTYr 1018
|||||
DB 2695 TCAGATGTGTGAGCTTGAGTGTCTATACGAACTTTTCACATAC 2742

RESULT 8
US-08-805-445-2
Sequence 2, Application US/08805445
Patent No. 5821069
GENERAL INFORMATION:
APPLICANT: Wilks, Andrew F.; Ziemiacki, Andrew;
APPLICANT: Harpur, Ailsa
TITLE OF INVENTION: NO. 5821069e1 Protein Tyrosine Kinase
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Felfe & Lynch,
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
COMPUTER: IBM PS/2

OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/805,445
FILING DATE: 25-FEB-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/446,038
FILING DATE: 19-MAY-1995
APPLICATION NUMBER: 08/064,067
FILING DATE: 30-Jun-1993
APPLICATION NUMBER: PCT/US91/08889
FILING DATE: 26-Jun-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: Australian PK3594/90
FILING DATE: 28-Jun-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: Australian 88229/91
FILING DATE: 27-Jun-1991
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 5821069man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 5244
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-688-9200
TELEFAX: 212-838-3884
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 3495 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: nucleic acid
US-08-805-445-2

Alignment Scores:
Pred. No.: 7e-05 Length: 3495
Score: 16.00 Matches: 16
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.46% Indels: 0
Gaps: 0
DB: 1

US-09-397-967-16 (1-1099) x US-08-805-445-2 (1-3495)

QY 1003 SeraspvAITPserPheglYvalLeuTYrgluLeuPheThrTYr 1018
|||||
DB 2695 TCAGATGTGTGAGCTTGAGTGTCTATACGAACTTTTCACATAC 2742

RESULT 9
US-08-064-067D-2
Sequence 2, Application US/08064067D
Patent No. 5852184
GENERAL INFORMATION:
APPLICANT: Wilks, Andrew F.; Ziemiacki, Andrew;
APPLICANT: Harpur, Ailsa
TITLE OF INVENTION: NO. 5852184e1 Protein Tyrosine Kinase
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Felfe & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/064,067D
FILING DATE: 30-Jun-1993

PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US91/08889
FILING DATE: 26-NOV-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: Australian PK3594/90
FILING DATE: 28-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: Australian 88229/91
FILING DATE: 27-NOV-1991
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 5852184man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 5244
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-688-9200
TELEFAX: 212-688-3884
INFORMATION FOR SEQ. ID NO.: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 3495 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: nucleic acid
US-08-064-067D-2

Alignment Scores:
Pred. No.: 7e-05 Length: 3495
Score: 16.00 Matches: 16
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.46% Indels: 0
DB: Gaps: 2

US-09-397-967-16 (1-1099) x US-08-064-067D-2 (1-3495)

QY 1003 serAspValTrpSerPheGlyValValLeuTyrgLuleuPheThrTyr 1018
DB 2695 TCAGATGTGGAGCTTGTGAGTGTCTATACGAACTTTTCACATAC 2742

RESULT 10
US-09-066-208-2
Sequence 2, Application US/09066208
Patent No. 5910426
GENERAL INFORMATION:
APPLICANT: Wilks, Andrew F.; Ziemlecki, Andrew;
APPLICANT: Harpur, Alisa
TITLE OF INVENTION: No. 5910426el Protein Tyrosine Kinase
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Felte & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/066,208
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/805,445
FILING DATE: 25-FEB-1997
APPLICATION NUMBER: US 08/446,038
FILING DATE: 19-MAY-1995
APPLICATION NUMBER: 08/064,067
FILING DATE: 30-JUN-1993
APPLICATION NUMBER: PCT/US91/08889
FILING DATE: 26-NOV-1991

PRIOR APPLICATION DATA:
APPLICATION NUMBER: Australian PK3594/90
FILING DATE: 28-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: Australian 88229/91
FILING DATE: 27-NOV-1991
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 5910426man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 5244
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-688-9200
TELEFAX: 212-688-3884
INFORMATION FOR SEQ. ID NO.: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 3495 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: nucleic acid
US-09-066-208-2

Alignment Scores:
Pred. No.: 7e-05 Length: 3495
Score: 16.00 Matches: 16
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.46% Indels: 0
DB: Gaps: 2

US-09-397-967-16 (1-1099) x US-09-066-208-2 (1-3495)

QY 1003 serAspValTrpSerPheGlyValValLeuTyrgLuleuPheThrTyr 1018
DB 2695 TCAGATGTGGAGCTTGTGAGTGTCTATACGAACTTTTCACATAC 2742

RESULT 11
US-08-980-080-3
Sequence 3, Application US/08980080
Patent No. 6312941
GENERAL INFORMATION:
APPLICANT: CARTER-SU, CHRISTIN
APPLICANT: ROI, LIANG-YOU
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING
TITLE OF INVENTION: SIGNALING PATHWAY AGONISTS AND ANTAGONISTS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: MEDLEN & CARROLL, LLP
STREET: 220 MONTGOMERY STREET, SUITE 2200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/980,080
FILING DATE: 26-NOV-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: CARROLL, PETER G.
REGISTRATION NUMBER: 32,837
REFERENCE/DOCKET NUMBER: UM-03036
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ. ID NO.: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3495 base pairs
TYPE: nucleic acid

STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..2979
US-08-980-080-3

Alignment Scores:
Pred. No.: 7e-05
Score: 16.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 1.46%
DB: 4
Gaps: 0

US-09-397-967-16 (1-1099) x US-08-980-080-3 (1-3495)

QY 1003 SeraspVAlTrpSerPheGlyValValLeuTyrgluLeuPheThrTYR 1018
|||||
Db 2695 TCAGATGTGTGAGCTTTGGAGTGTCTATACGACACTTTTCACATAC 2742

RESULT 12
US-08-097-997A-8
Sequence 8, Application US/08097997A
Patent No. 5728536
GENERAL INFORMATION:
APPLICANT: Ihle, James N.
APPLICANT: Silvenoinen, Olli
APPLICANT: Withuhn, Bruce A.
APPLICANT: Quelle, Frederick W.
TITLE OF INVENTION: Jak Kinases and Regulation of Cytokine Signal
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/097,997A
FILING DATE: 29-JUL-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Fox, Samuel L.
REGISTRATION NUMBER: 30,353
REFERENCE/DOCKET NUMBER: 0656,0370000/SLF/GKT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 3629 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 94..3480
US-08-097-997A-8

Alignment Scores:
Pred. No.: 7.25e-05
Score: 16.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 1.46%
DB: 0
Gaps: 0

Query Match: 1.46%
DB: 1
Indels: 0
Gaps: 0

US-09-397-967-16 (1-1099) x US-08-097-997A-8 (1-3629)

QY 1003 SeraspVAlTrpSerPheGlyValValLeuTyrgluLeuPheThrTYR 1018
|||||
Db 3196 TCAGATGTGTGAGCTTTGGAGTGTCTATACGACACTTTTCACATAC 3243

RESULT 13
US-08-665-574C-8
Sequence 8, Application US/08665574C
Patent No. 6136595
GENERAL INFORMATION:
APPLICANT: Ihle, James N.
APPLICANT: Silvenoinen, Olli
APPLICANT: Withuhn, Bruce A.
TITLE OF INVENTION: Jak Kinases and Regulation of Cytokine
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/665,574C
FILING DATE: 18-JUN-1996
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/282,012
FILING DATE: 29-JUL-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/097,997
FILING DATE: 29-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/118,968
FILING DATE: 09-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: Bugalsky, Lawrence B.
REGISTRATION NUMBER: 35,086
REFERENCE/DOCKET NUMBER: 0656,0370002/SLF/LBB
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 3629 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 94..3480
US-08-665-574C-8

Alignment Scores:
Pred. No.: 7.25e-05
Score: 16.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 1.46%
DB: 3
Indels: 0
Gaps: 0

US-09-397-967-16 (1-1099) x US-08-665-574C-8 (1-3629)

Oy 1003 SeraspValITrpSerPheGlyValValleuTyrgIuleuphetTyr 1018
|||||
Db 3196 TCAGATGTGTGAGCTTTGGAGTGTCTCTATACGACTTTTCACATAC 3243

RESULT 14

US-08-946-994-8

Sequence 8, Application US/08946994
Patent No. 6210654

GENERAL INFORMATION:

APPLICANT: Ihle, James N.

APPLICANT: Silvenoinen, Olli

APPLICANT: Wilthuhn, Bruce A.

TITLE OF INVENTION: Jak Kinases and Regulation of Cytokine Signal

NUMBER OF SEQUENCES: 17

CORRESPONDENCE ADDRESS:

ADDRESSEE: Sterne, Kessler, Goldstein & Fox

STREET: 1100 New York Avenue, Suite 600

CITY: Washington

STATE: D.C.

COUNTRY: U.S.A.

ZIP: 20005-3934

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/946,994

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/665,574

FILING DATE: 18-JUN-1996

APPLICATION NUMBER: 08/282,012

FILING DATE: 29-JUL-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/097,997

FILING DATE: 29-JUL-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/118,968

FILING DATE: 09-SEP-1993

ATTORNEY/AGENT INFORMATION:

NAME: Fox, Samuel L.

REGISTRATION NUMBER: 30,353

REFERENCE/DOCKET NUMBER: 0656.0370002/SLF/GKT

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 371-2540

TELEFAX: (202) 371-2540

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 3629 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

FEATURE:

NAME/KEY: CDS

LOCATION: 94..3480

US-08-946-994-8

Alignment Scores:

Pred. No.: 7.25e-05

Score: 16.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 1.468

DB: 4

Gaps: 0

US-09-397-967-16 (1-1099) x US-08-946-994-8 (1-3629)

Oy 1003 SeraspValITrpSerPheGlyValValleuTyrgIuleuphetTyr 1018

|||||

Db 3196 TCAGATGTGTGAGCTTTGGAGTGTCTCTATACGACTTTTCACATAC 3243

RESULT 15

US-08-567-508C-1

Sequence 1, Application US/08567508C
Patent No. 5914393

GENERAL INFORMATION:

APPLICANT: Coleman, Roger

APPLICANT: Stuart, Susan G.

TITLE OF INVENTION: A NOVEL HUMAN JAK2 KINASE

NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESS:

ADDRESSEE: INCYTE PHARMACEUTICALS, INC.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: US

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ Version 1.5

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/567,508C

FILING DATE: 05-DEC-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-004905

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650-845-4166

TELEFAX: 650-855-0555

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 4482 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

IMMEDIATE SOURCE:

LIBRARY: Placenta

CLONE: 179527

US-08-567-508C-1

Alignment Scores:

Pred. No.: 8.78e-05

Score: 16.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 1.468

DB: 2

Gaps: 0

US-09-397-967-16 (1-1099) x US-08-567-508C-1 (1-4482)

Oy 1003 SeraspValITrpSerPheGlyValValleuTyrgIuleuphetTyr 1018

Db 3442 TCAGATGTGTGAGCTTTGGAGTGTCTCTATACGACTTTTCACATAC 3489

|||||

RESULT 16

US-09-196-480-1

Sequence 1, Application US/09196480

Patent No. 6019966

GENERAL INFORMATION:

APPLICANT: Coleman, Roger

APPLICANT: Stuart, Susan G.

TITLE OF INVENTION: A NOVEL HUMAN JAK2 KINASE

NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESS:

ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 Porter Drive
City: Palo Alto
STATE: CA
COUNTRY: US
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/196,480
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/567,508
FILING DATE: 05-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-004905
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4482 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
IMMEDIATE SOURCE:
LIBRARY: Flaccenta
CLONE: 179527
US-09-196-480-1
Alignment Scores:
Pred. No.: 8.78e-05 Length: 4482
Score: 16.00 Matches: 16
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.46% Indels: 0
DB: 3 Gaps: 0
US-09-397-967-16 (1-1099) x US-09-196-480-1 (1-4482)
Oy 1003 SeraspValtIrpSerPheGlyValLeuTYRGluLeuPheThTyr 1018
|||||
Db 3442 TCAGATGTTTGAGCTTGAGTGGTGTGATGAACCTTTCACATAC 3489
RESULT 17
US-08-481-003-9/C
Sequence 9, Application US/08481003
GENERAL INFORMATION:
APPLICANT: CAPON, DANIEL J
APPLICANT: TIAN, HUAN
APPLICANT: SMITH, DOUGLAS H
APPLICANT: WINSLOW, GENINE A
APPLICANT: SIEKEVITZ, MIRIAM
TITLE OF INVENTION: CHIMERIC RECEPTORS FOR REGULATING
TITLE OF INVENTION: CELLULAR PROLIFERATION AND EFFECTOR FUNCTION
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: CELL GENESYS, INC.
STREET: 322 LAKESIDE DRIVE
CITY: FOSTER CITY
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/481,003
FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/382,846
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: KRUPEN, KAREN I
REGISTRATION NUMBER: 34,647
REFERENCE/DOCKET NUMBER: CELL 17
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 358-9600 x131
TELEFAX: (415) 349-7392
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 41 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-481-003-9
Alignment Scores:
Pred. No.: 0.000652 Length: 41
Score: 13.00 Matches: 13
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.18% Indels: 0
DB: 1 Gaps: 0
US-09-397-967-16 (1-1099) x US-08-481-003-9 (1-41)
Oy 1 MetalapProserGluGluThrProleuIleProGln 13
|||||
Db 40 ATGGACCTCCAGTGAAGAGACACCTCTGATCCCTCAG 2
RESULT 18
US-08-485-598-9/C
Sequence 9, Application US/08485598
Patent No. 6077947
GENERAL INFORMATION:
APPLICANT: CAPON, DANIEL J
APPLICANT: TIAN, HUAN
APPLICANT: SMITH, DOUGLAS H
APPLICANT: WINSLOW, GENINE A
APPLICANT: SIEKEVITZ, MIRIAM
TITLE OF INVENTION: CHIMERIC RECEPTORS FOR REGULATING
TITLE OF INVENTION: CELLULAR PROLIFERATION AND EFFECTOR FUNCTION
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: CELL GENESYS, INC.
STREET: 322 LAKESIDE DRIVE
CITY: FOSTER CITY
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,598
FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/382,846
FILING DATE:
ATTORNEY/AGENT INFORMATION:

NAME: KRUPEN, KAREN I
REGISTRATION NUMBER: 34,647
REFERENCE/DOCKET NUMBER: CELL 17
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 358-9600 x131
TELEFAX: (415) 349-7392
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 41 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-485-598-9

Alignment Scores:
Pred. No.: 0.000652 Length: 41
Score: 13.00 Matches: 13
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.18% Indels: 0
DB: 3 Gaps: 0

US-09-397-967-16 (1-1099) x US-08-485-598-9 (1-41)

OY 1 MetAlaProSerGluThrProLeuIleProGln 13
Db 40 ATGGCACCTCCAACTGAGAGACACCTGATCCCTCAG 2

RESULT 19
US-08-481-003-8

Sequence 8, Application US/08481003
Patent No. 5741899

GENERAL INFORMATION:

APPLICANT: CAPON, DANIEL J
APPLICANT: TIAN, HUAN

APPLICANT: SMITH, DOUGLAS H
APPLICANT: WINSLOW, GENINE A

APPLICANT: SIEKEVITZ, MIRIAM
TITLE OF INVENTION: CHIMERIC RECEPTORS FOR REGULATING

TITLE OF INVENTION: CELLULAR PROLIFERATION AND EFFECTOR FUNCTION
NUMBER OF SEQUENCES: 31

CORRESPONDENCE ADDRESS:
ADDRESSEE: CELL GENESYS, INC.

STREET: 322 LAKESIDE DRIVE
CITY: FOSTER CITY

STATE: CALIFORNIA
COUNTRY: USA

ZIP: 94404
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/481,003

FILING DATE:
CLASSIFICATION: 536

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/382,846

FILING DATE:
ATTORNEY/AGENT INFORMATION:

NAME: KRUPEN, KAREN I
REGISTRATION NUMBER: 34,647

REFERENCE/DOCKET NUMBER: CELL 17
TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 358-9600 x131
TELEFAX: (415) 349-7392

INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:

LENGTH: 45 base pairs
TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

US-08-485-598-8

TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-481-003-8

Alignment Scores:
Pred. No.: 0.00071 Length: 45
Score: 13.00 Matches: 13
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.18% Indels: 0
DB: 1 Gaps: 0

US-09-397-967-16 (1-1099) x US-08-481-003-8 (1-45)

OY 1 MetAlaProSerGluThrProLeuIleProGln 13
Db 6 ATGGCACCTCCAACTGAGAGACACCTGATCCCTCAG 44

RESULT 20
US-08-485-598-8

Sequence 8, Application US/08485598
Patent No. 6077947

GENERAL INFORMATION:

APPLICANT: CAPON, DANIEL J
APPLICANT: TIAN, HUAN

APPLICANT: SMITH, DOUGLAS H
APPLICANT: WINSLOW, GENINE A

APPLICANT: SIEKEVITZ, MIRIAM
TITLE OF INVENTION: CHIMERIC RECEPTORS FOR REGULATING

TITLE OF INVENTION: CELLULAR PROLIFERATION AND EFFECTOR FUNCTION
NUMBER OF SEQUENCES: 31

CORRESPONDENCE ADDRESS:
ADDRESSEE: CELL GENESYS, INC.

STREET: 322 LAKESIDE DRIVE
CITY: FOSTER CITY

STATE: CALIFORNIA
COUNTRY: USA

ZIP: 94404
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,598

FILING DATE:
CLASSIFICATION: 536

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/382,846

FILING DATE:
ATTORNEY/AGENT INFORMATION:

NAME: KRUPEN, KAREN I
REGISTRATION NUMBER: 34,647

REFERENCE/DOCKET NUMBER: CELL 17
TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 358-9600 x131
TELEFAX: (415) 349-7392

INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:

LENGTH: 45 base pairs
TYPE: nucleic acid

STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

US-08-485-598-8

Alignment Scores:

Pred. No.: 0.00071 Length: 45
Score: 13.00 Matches: 13
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.18% Indels: 0
DB: 3 Gaps: 0

APPLICANT: Wood, William I.
TITLE OF INVENTION: PROTEIN TYROSINE KINASE AGONIST ANTIBODIES
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Palin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/04228
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/222616
FILING DATE: 04-APR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Wendy M. Lee
REGISTRATION NUMBER: 00,000
REFERENCE/DOCKET NUMBER: 821P3PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1994
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 151 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
PCT-US95-04228-11

Alignment Scores:
Pred. No.: 0 0173 Length: 151
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 5 Gaps: 0

US-09-397-967-16 (1-1099) x PCT-US95-04228-11 (1-151)
QY 941 VALHISARGASPLEUALAALARGASNIILEUVAL 952
DB 1 GTGCACAGGAGATCTCGCGCTCGAACAATCTGCTC 36

RESULT 24
US-08-604-989A-8
Sequence 8, Application US/08604989A
Patent No. 5834208
GENERAL INFORMATION:
APPLICANT: Sakano, S.
TITLE OF INVENTION: No. 5834208el Tyrosine Kinase
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/604,989A
FILING DATE: February 23, 1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Charles E. Miller
REGISTRATION NUMBER: 24,576
REFERENCE/DOCKET NUMBER: 1920-026
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 738 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
ORIGINAL SOURCE:
ORGANISM: human
STRAIN: UT-7
US-08-604-989A-8

Alignment Scores:
Pred. No.: 0.0732 Length: 738
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 2 Gaps: 0

US-09-397-967-16 (1-1099) x US-08-604-989A-8 (1-738)
QY 941 VALHISARGASPLEUALAALARGASNIILEUVAL 952
DB 349 GTGCACGCGAGCTGCGCGCCGCAACAATCTGCTC 384

RESULT 25
US-08-604-989A-9
Sequence 9, Application US/08604989A
Patent No. 5834208
GENERAL INFORMATION:
APPLICANT: Sakano, S.
TITLE OF INVENTION: No. 5834208el Tyrosine Kinase
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/604,989A
FILING DATE: February 23, 1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Charles E. Miller
REGISTRATION NUMBER: 24,576
REFERENCE/DOCKET NUMBER: 1920-026
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 1398 base pairs
TYPE: nucleic acid

```
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
ORIGINAL SOURCE:
ORGANISM: human
STRAIN: UT-7
US-08-604-989A-9

Alignment Scores:
Pred. No.: 0.131 length: 1398
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 2 Gaps: 0

US-09-397-967-16 (1-1099) x US-08-604-989A-9 (1-1398)
OY 941 VALHISARGASPLEUALAALAARGASNILEUVAL 952
DB 922 GTGCACCGCGACCTGCGCCGCCGCAACATCTGCTC 957

RESULT 26
US-08-604-989A-10
Sequence 10, Application US/08604989A
Patent No. 5834208
GENERAL INFORMATION:
APPLICANT: Sakano, S.
TITLE OF INVENTION: No. 5834208el Tyrosine Kinase
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/604,989A
FILING DATE: February 23, 1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Charles E. Miller
REGISTRATION NUMBER: 24,576
REFERENCE/DOCKET NUMBER: 1920-026
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 1521 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
ORIGINAL SOURCE:
ORGANISM: human
STRAIN: UT-7
US-08-604-989A-10

Alignment Scores:
Pred. No.: 0.141 length: 1521
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 2 Gaps: 0

STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
ORIGINAL SOURCE:
ORGANISM: human
STRAIN: UT-7
US-08-604-989A-10

Alignment Scores:
Pred. No.: 0.157 length: 1713
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 4 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-741-154-1 (1-1713)
OY 941 VALHISARGASPLEUALAALAARGASNILEUVAL 952
DB 991 GTGCACCGCGACCTGCGCCGCCGCAACATCTGCTC 1026

RESULT 28
US-08-604-989A-11
Sequence 11, Application US/08604989A
Patent No. 5834208
GENERAL INFORMATION:
APPLICANT: Sakano, S.
TITLE OF INVENTION: No. 5834208el Tyrosine Kinase
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/604,989A
FILING DATE: February 23, 1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Charles E. Miller
REGISTRATION NUMBER: 24,576
REFERENCE/DOCKET NUMBER: 1920-026
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
```

;; INFORMATION FOR SEQ ID NO: 11:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1942 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: linear
;; MOLECULE TYPE: cDNA to mRNA
;; ORIGINAL SOURCE:
;; ORGANISM: human
;; STRAIN: UT-7
US-08-604-989A-11

Alignment Scores:

Pred. No.:	0.177	Length:	1942
Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.09%	Indels:	0
DB:	2	Gaps:	0

US-09-397-967-16 (1-1099) x US-08-604-989A-11 (1-1942)

QY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952

Db 1307 GTGCACCGCGACCTGGCCGCCGCAACATCTGTGTC 1342

RESULT 30

US-08-876-882-1

Sequence 1, Application US/08876882

Patent No. 5981201

GENERAL INFORMATION:

APPLICANT: Avraham, Hava

TITLE OF INVENTION: METHODS OF DETECTION AND TREATMENT

TITLE OF INVENTION: OF BREAST CANCER

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:

ADDRESSEE: Hamilton, Brook, Smith & Reynolds P.C.

STREET: Two Militta Drive

CITY: Lexington

STATE: MA

COUNTRY: USA

ZIP: 02173-4799

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: Windows

SOFTWARE: FastSeq for Windows Version 2.0b

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/876,882

FILING DATE: 16-JUN-1997

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/035,228

ATTORNEY/AGENT INFORMATION:

NAME: Doreen, Hogle M

REGISTRATION NUMBER: 36,361

REFERENCE/DOCKET NUMBER: NEDH97-01PA

TELECOMMUNICATION INFORMATION:

TELEPHONE: 781-861-6240

TELEFAX: 781-861-9540

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1987 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

US-08-876-882-1

Alignment Scores:

Pred. No.: 0.18

Length: 1987

Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.09%	Indels:	0
DB:	2	Gaps:	0

US-09-397-967-16 (1-1099) x US-08-876-882-1 (1-1987)

QY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952

Db 1307 GTGCACCGCGACCTGGCCGCCGCAACATCTGTGTC 1342

RESULT 30

US-09-315-928-1

Sequence 1, Application US/09315928

Patent No. 6368796

GENERAL INFORMATION:

APPLICANT: Avraham, Hava

TITLE OF INVENTION: METHODS OF DETECTION AND TREATMENT OF

TITLE OF INVENTION: BREAST CANCER

FILE REFERENCE: NEDH97-01PAZ

CURRENT APPLICATION NUMBER: US/09/315,928

CURRENT FILING DATE: 1999-05-20

PRIOR APPLICATION NUMBER: US 08/876,882

PRIOR FILING DATE: 1997-06-16

PRIOR APPLICATION NUMBER: US 60/035,228

PRIOR FILING DATE: 1997-01-08

NUMBER OF SEQ ID NOS: 5

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 1

LENGTH: 1987

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: (263)...(1846)

US-09-315-928-1

Alignment Scores:

Pred. No.:	0.18	Length:	1987
Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.09%	Indels:	0
DB:	4	Gaps:	0

US-09-397-967-16 (1-1099) x US-09-315-928-1 (1-1987)

QY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952

Db 1307 GTGCACCGCGACCTGGCCGCCGCAACATCTGTGTC 1342

RESULT 31

US-08-426-509A-1

Sequence 1, Application US/08426509A

Patent No. 6326469

GENERAL INFORMATION:

APPLICANT: Ullrich, Axel

APPLICANT: Gishizky, Mikhail

TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN

TITLE OF INVENTION: TYROSINE KINASES

NUMBER OF SEQUENCES: 21

CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds

STREET: 1155 Avenue of the Americas

CITY: New York,

STATE: NY

COUNTRY: USA

ZIP: 10036-2711

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTED Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/426,509A
FILING DATE: 21-APR-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/232,545
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7683-0074-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-790-9090
TELEFAX: 212-869-9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2000 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
US-08-426-509A-1

Alignment Scores:
Pred. No.: 0.181 Length: 2000
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
Gaps: 0
DB: 4

US-09-397-967-16 (1-1099) x US-08-426-509A-1 (1-2000)

QY 941 ValHisArgAspLeuAlaIaArgAsnIleLeuVal 952
Db 1302 GTGCACCGGACCTGGCCGCCGACACATCTGTGTC 1337

RESULT 32
PCT-US95-05008-1
Sequence 1, Application PC/TUS9505008
GENERAL INFORMATION:
APPLICANT: Sugen, Inc.
APPLICANT: 515 Galveston Drive
APPLICANT: Redwood City, California 94063-4720
APPLICANT: United States of America
APPLICANT: Missenschaften E.V.
APPLICANT: Hofgarten Str. 2
APPLICANT: Munchen 80539
APPLICANT: Germany
TITLE OF INVENTION: Novel Megakaryocytic Protein Tyrosine
TITLE OF INVENTION: Kinases
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/05008
FILING DATE: 24-APR-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/232,545

FILING DATE: 22-APR-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7683-0074
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)790-9090
TELEFAX: (212)869-9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2000 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA
PCT-US95-05008-1

Alignment Scores:
Pred. No.: 0.181 Length: 2000
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
Gaps: 0
DB: 5

US-09-397-967-16 (1-1099) x PCT-US95-05008-1 (1-2000)

QY 941 ValHisArgAspLeuAlaIaArgAsnIleLeuVal 952
Db 1302 GTGCACCGGACCTGGCCGCCGACACATCTGTGTC 1337

RESULT 33
US-08-162-809-5
Sequence 5, Application US/08162809
Patent No. 5457048
GENERAL INFORMATION:
APPLICANT: Pasquale, Elena B.
APPLICANT: Sajjad, Pereydon G.
TITLE OF INVENTION: NOVEL EPH-RELATED TYROSINE KINASES,
TITLE OF INVENTION: NUCLEOTIDE SEQUENCES, AND METHODS OF USE
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: CAMPBELL AND FLORES
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: United States of America
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/162,809
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LJ 9503
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 2820 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
FEATURE:

```

; NAME/KEY: CDS
; LOCATION: 2..2548
US-08-162-809-5

Alignment Scores:
Pred. No.: 0.248 Length: 2820
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 1 Gaps: 0

US-09-397-967-16 (1-1099) x US-08-162-809-5 (1-2820)
OY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
Db 1817 GTGCATCGGAGTCTAGCTGCTCGAACATATCTGTC 1852

RESULT 34
US-08-449-645A-10
; Sequence 10, Application US/08449645A
; Patent No. 5981245
; GENERAL INFORMATION:
; APPLICANT: Fox, Gary M.
; TITLE OF INVENTION: Eph-Like Receptor Protein Tyrosine
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Patent Operations/RBW
; STREET: 1840 Dehavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/449,645A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Winter, Robert B.
; REFERENCE/DOCKET NUMBER: A-287
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2962 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2913
US-08-449-645A-10

Alignment Scores:
Pred. No.: 0.259 Length: 2962
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 2 Gaps: 0

US-09-397-967-16 (1-1099) x US-08-449-645A-10 (1-2962)
OY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
Db 2179 GTTCACCGTGAAGCTGGCTGCCCGCAACATCTCTGTC 2214

RESULT 35
US-08-702-367A-10
; Sequence 10, Application US/08702367A
; Patent No. 5981246
; GENERAL INFORMATION:
; APPLICANT: Fox, Gary M.
; TITLE OF INVENTION: Eph-Like Receptor Protein Tyrosine
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Patent Operations/RBW
; STREET: 1840 Dehavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/702,367A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Winter, Robert B.
; REFERENCE/DOCKET NUMBER: A-287
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2962 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2913
US-08-702-367A-10

Alignment Scores:
Pred. No.: 0.259 Length: 2962
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 2 Gaps: 0

US-09-397-967-16 (1-1099) x US-08-702-367A-10 (1-2962)
OY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
Db 2179 GTTCACCGTGAAGCTGGCTGCCCGCAACATCTCTGTC 2214

RESULT 36
PCT-US95-04681-10
; Sequence 10, Application PC/TUS9504681
; GENERAL INFORMATION:
; APPLICANT: Fox, Gary M.
; TITLE OF INVENTION: Eph-Like Receptor Protein Tyrosine
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Patent Operations/RBW
; STREET: 1840 Dehavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
```

;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: PCT/US95/04681
;; FILING DATE:
;; CLASSIFICATION:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Winter, Robert B.
;; REFERENCE/DOCKET NUMBER: A-287
;; INFORMATION FOR SEQ ID NO: 10:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 2962 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: cDNA
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: 1..2913
;; PCT-US95-04681-10

Alignment Scores:
Pred. No.: 0.259 Length: 2962
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: Gaps: 0

US-09-397-967-16 (1-1099) x PCT-US95-04681-10 (1-2962)

QY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
Db 2179 GTTCACCGTGACCTGCGCCGCCACATCCTGTC 2214
|||||

RESULT 37
US-08-348-143-2
; Sequence 2, Application US/08348143
; Patent No. 5506205
; GENERAL INFORMATION:
; APPLICANT: TAJIMA, HISAO
; APPLICANT: KITAGAWA, KOICHIRO
; APPLICANT: OHNO, HIROYUKI
; TITLE OF INVENTION: A No. 5506205el Polypeptide of protein p140 and DNAs
; TITLE OF INVENTION: encoding it
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SUCHRUE, MION, ZINN, MACPEAK & SEAS
; STREET: 2100 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20037-3202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/348,143
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 315806/1993
; FILING DATE: 24-NOV-1993
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)293-7060
; TELEFAX: (202)293-7860
; TELEX: 6491103
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2982 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single

;; TOPOLOGY: linear
;; MOLECULE TYPE: cDNA to mRNA
;; ORIGINAL SOURCE:
;; ORGANISM: rat
;; TISSUE TYPE: skeletal muscle myoblast
;; CELL LINE: L6
;; US-08-348-143-2

Alignment Scores:
Pred. No.: 0.261 Length: 2982
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: Gaps: 0

US-09-397-967-16 (1-1099) x US-08-348-143-2 (1-2982)

QY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
Db 2248 GTGCACCGTGACCTGCGCCGCCACATCCTGTC 2283
|||||

RESULT 38
US-08-571-785-2
; Sequence 2, Application US/08571785
; Patent No. 5804411
; GENERAL INFORMATION:
; APPLICANT: TAJIMA, HISAO
; APPLICANT: KITAGAWA, KOICHIRO
; APPLICANT: OHNO, HIROYUKI
; TITLE OF INVENTION: A No. 5804411el Polypeptide of protein p140
; TITLE OF INVENTION: and DNAs encoding it
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SUCHRUE, MION, ZINN, MACPEAK & SEAS
; STREET: 2100 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20037-3202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/571,785
; FILING DATE: 13-DEC-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/348,143
; FILING DATE: 23-NOV-1994
; APPLICATION NUMBER: JP 315806/1993
; FILING DATE: 24-NOV-1993
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)293-7060
; TELEFAX: (202)293-7860
; TELEX: 6491103
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2982 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; ORIGINAL SOURCE:
; ORGANISM: rat
; TISSUE TYPE: skeletal muscle myoblast
; CELL LINE: L6
;; US-08-571-785-2

Alignment Scores:

Pred. No.: 0.261 Length: 2982
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 1 Gaps: 0

US-09-397-967-16 (1-1099) x US-08-571-785-2 (1-2982)

OY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952

DB 2248 GTGCACCGTGACCTGCTGCGCCGCAACATCTTGTG 2283

RESULT 39

US-09-192-435-2

Sequence 2, Application US/09192435

Patent No. 6303320

GENERAL INFORMATION:

APPLICANT: TAJIMA, HISAO

APPLICANT: KITAGAWA, KOICHIRO

APPLICANT: OHNO, HIROYUKI

APPLICANT: UENO, TOSHIO

TITLE OF INVENTION: A No. 6303320el Polypeptide of Protein p140

NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS: 16

ADDRESS: SUGHRUE, MION, ZINN, MACPEAK & SEAS

STREET: 2100 Pennsylvania Avenue, N.W.

CITY: Washington

STATE: D.C.

COUNTRY: U.S.A.

ZIP: 20037-3202

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/192,435

FILING DATE:

CLASSIFICATION:

Prior Application Data:

APPLICATION NUMBER: 08/571,785

FILING DATE: 13-DEC-1995

APPLICATION NUMBER: 08/348,143

FILING DATE: 23-NOV-1994

APPLICATION NUMBER: JP 315806/1993

FILING DATE: 24-NOV-1993

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202)293-7060

TELEFAX: (202)293-7860

TELEX: 6491103

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 2982 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA to mRNA

ORIGINAL SOURCE:

ORGANISM: rat

TISSUE TYPE: skeletal muscle myoblast

CELL LINE: L6

US-09-192-435-2

Alignment Scores:

Pred. No.: 0.261 Length: 2982

Score: 12.00 Matches: 12

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 1.09% Indels: 0

DB: 4 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-192-435-2 (1-2982)

OY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952

DB 2248 GTGCACCGTGACCTGCTGCGCCGCAACATCTTGTG 2283

RESULT 40

US-09-558-340-2

Sequence 2, Application US/09558340

Patent No. 6432913

GENERAL INFORMATION:

APPLICANT: TAJIMA, HISAO

APPLICANT: KITAGAWA, KOICHIRO

APPLICANT: OHNO, HIROYUKI

APPLICANT: UENO, TOSHIO

TITLE OF INVENTION: A No. 6432913el Polypeptide of Protein p140

NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS: 16

ADDRESS: SUGHRUE, MION, ZINN, MACPEAK & SEAS

STREET: 2100 Pennsylvania Avenue, N.W.

CITY: Washington

STATE: D.C.

COUNTRY: U.S.A.

ZIP: 20037-3202

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/558,340

FILING DATE: 26-APR-2000

CLASSIFICATION:

Prior Application Data:

APPLICATION NUMBER: 09/192,435

FILING DATE: 08-JAN-1998

APPLICATION NUMBER: 08/571,785

FILING DATE: 13-DEC-1995

APPLICATION NUMBER: 08/348,143

FILING DATE: 23-NOV-1994

APPLICATION NUMBER: JP 315806/1993

FILING DATE: 24-NOV-1993

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202)293-7060

TELEFAX: (202)293-7860

TELEX: 6491103

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 2982 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA to mRNA

ORIGINAL SOURCE:

ORGANISM: rat

TISSUE TYPE: skeletal muscle myoblast

CELL LINE: L6

US-09-558-340-2

Alignment Scores:

Pred. No.: 0.261 Length: 2982

Score: 12.00 Matches: 12

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 1.09% Indels: 0

DB: 4 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-558-340-2 (1-2982)

OY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952

DB 2248 GTGCACCGTGACCTGCTGCGCCGCAACATCTTGTG 2283

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Winter, Robert B.
REFERENCE/DOCKET NUMBER: A-287
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 3116 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 34..2994
US-08-702-367A-14

Alignment Scores:
Pred. No.: 0.271 Length: 3116
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 2 Gaps: 0

US-09-397-967-16 (1-1099) x US-08-702-367A-14 (1-3116)

OY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
Db 2260 GTCCATCGTATCTGCGCCAGCAACATCTGCTG 2295

RESULT 44
PCT-US95-04681-14
Sequence 14, Application PC/TUS9504681
GENERAL INFORMATION:
APPLICANT: Fox, Gary M.
TITLE OF INVENTION: EPH-Like Receptor Protein Tyrosine
NUMBER OF INVENTION: Kinases
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amgen Patent Operations/RBW
STREET: 1840 Delavilland Drive
CITY: Thousand Oaks
STATE: California
COUNTRY: USA
ZIP: 91320

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/04681
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Winter, Robert B.
REFERENCE/DOCKET NUMBER: A-287
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 3116 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 34..2994
PCT-US95-04681-14

Alignment Scores:
Pred. No.: 0.271 Length: 3116
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 2 Gaps: 0

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 5 Gaps: 0

US-09-397-967-16 (1-1099) x PCT-US95-04681-14 (1-3116)

OY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
Db 2260 GTCCATCGTATCTGCGCCAGCAACATCTGCTG 2295

RESULT 45
US-08-162-809-1
Sequence 1, Application US/08162809
Patent No. 5457048
GENERAL INFORMATION:
APPLICANT: Pasquale, Elena B.
APPLICANT: Sajjadi, Fereydoun G.
TITLE OF INVENTION: NOVEL EPH-RELATED TYROSINE KINASES,
NUMBER OF INVENTION: NUCLEOTIDE SEQUENCES, AND METHODS OF USE
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: CAMPBELL AND FLORES
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: United States of America
ZIP: 92122

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/162,809
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LJ 9503
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3133 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: John(3..419, 421..2858)
US-08-162-809-1

Alignment Scores:
Pred. No.: 0.273 Length: 3133
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 1 Gaps: 0

US-09-397-967-16 (1-1099) x US-08-162-809-1 (1-3133)

OY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
Db 2125 GTCCACAGGATCTGCGCCAGCAACATCTGCTG 2160

RESULT 46
US-09-166-350-23
Sequence 23, Application US/09166350A
Patent No. 6440663
GENERAL INFORMATION:

APPLICANT: Scanlan, Mathew
APPLICANT: Chen, Yao
APPLICANT: Stockert, Elisabeth
APPLICANT: Old, Lloyd
APPLICANT: Jager, Elke
APPLICANT: Knuth, Alex
TITLE OF INVENTION: Renal Cancer Associated Antigens and
FILE REFERENCE: 10461/7051
CURRENT APPLICATION NUMBER: US/09/166,350A
EARLIER FILING DATE: 1998-10-05
NUMBER OF SEQ ID NOS: 35
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO: 23
LENGTH: 3150
TYPE: DNA
ORGANISM: Homo sapiens
US-09-166-350-23

Alignment Scores:
Pred. No.: 0.274 Length: 3150
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 4 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-166-350-23 (1-3150)

QY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
DB 2236 GTTCACCGTGACCTGCTGCCGACATCTCTGTC 2271

RESULT 47
US-08-222-616-34

Sequence 34, Application US/08222616

Patent No. 5635177
GENERAL INFORMATION:
APPLICANT: Bennett, Brian D.
APPLICANT: Goeddel, David
APPLICANT: Lee, James M.
APPLICANT: Matthews, William
APPLICANT: Tsai, Shao Ping
APPLICANT: Wood, William I.
TITLE OF INVENTION: PROTEIN TYROSINE KINASE AGONIST
TITLE OF INVENTION: ANTIBODIES
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080

COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/222,616
FILING DATE: 4-APR-1994
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/00586
FILING DATE: 22-JAN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/826935
FILING DATE: 22-JAN-1992
ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M.

REGISTRATION NUMBER:
REFERENCE/DOCKET NUMBER: 821P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1994
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 3348 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-222-616-34

Alignment Scores:
Pred. No.: 0.29 Length: 3348
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 1 Gaps: 0

US-09-397-967-16 (1-1099) x US-08-222-616-34 (1-3348)

QY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
DB 2227 GTGCATCGTGATCTGCCGACGACATCTCTGTC 2262

RESULT 48
US-08-446-648-34

Sequence 34, Application US/08446648

Patent No. 6331302
GENERAL INFORMATION:
APPLICANT: Genentech, Inc.
APPLICANT: Bennett, Brian D.
APPLICANT: Goeddel, David
APPLICANT: Lee, James M.
APPLICANT: Matthews, William
APPLICANT: Tsai, Shao Ping
APPLICANT: Wood, William I.
TITLE OF INVENTION: PROTEIN TYROSINE KINASE AGONIST ANTIBODIES
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Winpatin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/446,648
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/222616
FILING DATE: 04-APR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M.
REGISTRATION NUMBER: 40,378
REFERENCE/DOCKET NUMBER: P0821P3PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1994
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 3348 base pairs
TYPE: Nucleic Acid

```

STRANDEDNESS: Single
TOPOLOGY: Linear
US-08-446-648-34

Alignment Scores:
Pred. No.: 0.29
Score: 12.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 1.09%
DB: 4

Length: 3348
Matches: 12
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0

US-09-397-967-16 (1-1099) x US-08-446-648-34 (1-3348)
QY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
DB 2227 GTGCATCGTGATCTGGCCGACGAGACATCTCTGCTG 2262

RESULT 49
PCT-US95-04228-34
Sequence 34, Application PC/TUS9504228
GENERAL INFORMATION:
APPLICANT: Genentech, Inc.
APPLICANT: Bennett, Brian D.
APPLICANT: Goeddel, David
APPLICANT: Lee, James M.
APPLICANT: Matthews, William
APPLICANT: Tsai, Siao Ping
APPLICANT: Wood, William I.
TITLE OF INVENTION: PROTEIN TYROSINE KINASE AGONIST ANTIBODIES
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: PC-DOS/MS-DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/04228
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/222616
FILING DATE: 04-APR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Wendy M. Lee
REGISTRATION NUMBER: 00,000
REFERENCE/DOCKET NUMBER: 821P3PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1994
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 3348 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
PCT-US95-04228-34

Alignment Scores:
Pred. No.: 0.29
Score: 12.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 1.09%
DB: 5

Length: 3348
Matches: 12
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0

```

```

US-09-397-967-16 (1-1099) x PCT-US95-04228-34 (1-3348)
QY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
DB 2227 GTGCATCGTGATCTGGCCGACGAGACATCTCTGCTG 2262

RESULT 50
US-08-162-809-9
Sequence 9, Application US/08162809
Patent No. 5457048
GENERAL INFORMATION:
APPLICANT: Pasquale, Elena B.
APPLICANT: Sajjadi, Fereydon G.
TITLE OF INVENTION: NOVEL EPH-RELATED TYROSINE KINASES,
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: CAMPBELL AND FLORES
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: United States of America
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/162,809
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LJ 9503
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-8949
TELEFAX: (619) 535-9001
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 3546 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 2..2920
US-08-162-809-9

Alignment Scores:
Pred. No.: 0.305
Score: 12.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 1.09%
DB: 1

Length: 3546
Matches: 12
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0

US-09-397-967-16 (1-1099) x US-08-162-809-9 (1-3546)
QY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
DB 2189 GTGCACCGAGACCTGGCTGCCGACACATCTCTGCTC 2224

RESULT 51
US-08-162-809-13
Sequence 13, Application US/08162809
Patent No. 5457048
GENERAL INFORMATION:
APPLICANT: Pasquale, Elena B.
APPLICANT: Sajjadi, Fereydon G.
TITLE OF INVENTION: NOVEL EPH-RELATED TYROSINE KINASES,

```

TITLE OF INVENTION: NUCLEOTIDE SEQUENCES, AND METHODS OF USE
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESSES:
ADDRESSEE: CAMPBELL AND FLORES
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: United States of America
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/162,809
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LJ 9503
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 3591 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 2..2965
US-08-162-809-13
Alignment Scores:
Pred. No.: 0.309 Length: 3591
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
Gaps: 0
DB: 1
US-09-397-967-16 (1-1099) x US-08-162-809-13 (1-3591)
QY 941 ValHisArgAspLeuAlaIaIaArgAsnIleLeuVal 952
Db 2234 GTGCACGAGACCTGCTGCCCGCACATCCTGCTC 2269
RESULT 52
US-08-469-537A-100
Sequence 100, Application US/08469537A
Patent No. 5843749
GENERAL INFORMATION:
APPLICANT: Maisompierre, et al.
TITLE OF INVENTION: EHK AND FOR TYROSINE
NUMBER OF SEQUENCES: 107
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Regeneron Pharmaceuticals, Inc.
STREET: 777 Old Saw Mill River Road
CITY: Tarrytown
STATE: NY
COUNTRY: U.S.A.
ZIP: 10591
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/469,537A

FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 08/406,247
FILING DATE: 17-MAR-1995
APPLICATION NUMBER: USSN 08/144,992
FILING DATE: 28-OCT-1993
APPLICATION NUMBER: USSN 07/736,559
FILING DATE: 26-JUL-1991
ATTORNEY/AGENT INFORMATION:
NAME: Kempler, Ph.D., Gall M
REGISTRATION NUMBER: 32,143
REFERENCE/DOCKET NUMBER: REG 070C
TELEPHONE: 914-345-7400
TELEFAX: 914-345-7721
TELEX:
INFORMATION FOR SEQ ID NO: 100:
SEQUENCE CHARACTERISTICS:
LENGTH: 3592 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA
FEATURE:
NAME/KEY: CDS
LOCATION: 598..3444
FEATURE:
NAME/KEY: modified_base
LOCATION: 56
OTHER INFORMATION: /mod_base= OTHER
OTHER INFORMATION: /label= N
OTHER INFORMATION: /note= "where N = G, A, C or T"
FEATURE:
NAME/KEY: modified_base
LOCATION: 3538
OTHER INFORMATION: /mod_base= OTHER
OTHER INFORMATION: /label= N
OTHER INFORMATION: /note= "where N = G, A, C or T"
US-08-469-537A-100
Alignment Scores:
Pred. No.: 0.309 Length: 3592
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
Gaps: 0
DB: 2
US-09-397-967-16 (1-1099) x US-08-469-537A-100 (1-3592)
QY 941 ValHisArgAspLeuAlaIaIaArgAsnIleLeuVal 952
Db 2977 GTTCATCGAGACCTAGACCTAGACATCTTGATG 3012
RESULT 53
US-08-162-809-7
Sequence 7, Application US/08162809
Patent No. 5457048
GENERAL INFORMATION:
APPLICANT: Pasquale, Elena B.
TITLE OF INVENTION: NOVEL EPH-RELATED TYROSINE KINASES,
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESSES:
ADDRESSEE: CAMPBELL AND FLORES
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: United States of America
ZIP: 92122
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA: #1.0, Version #1.25
APPLICATION NUMBER: US/08/162,809
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LJ 9503
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 3776 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 290..3208
US-08-162-809-7
Alignment Scores:
Pred. No.: 0.323 Length: 3776
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 1 Gaps: 0
US-09-397-967-16 (1-1099) x US-08-162-809-7 (1-3776)
QY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
Db 2477 GTGCATCGTGTCTCGACGACGATCACTTAGTC 2512
RESULT 54
US-08-436-044-5
Sequence 5, Application US/08436044
Patent No. 5624899
GENERAL INFORMATION:
APPLICANT: Bennett, Brian D.
APPLICANT: Matthews, William
TITLE OF INVENTION: HTK LIGAND
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/436,044
FILING DATE: 05-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/277722
FILING DATE: 20-JUL-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M.
REGISTRATION NUMBER: 00,000
REFERENCE/DOCKET NUMBER: 90203
TELECOMMUNICATION INFORMATION:

TELEPHONE: 415/225-1994
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 3969 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-436-044-5
Alignment Scores:
Pred. No.: 0.338 Length: 3969
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 1 Gaps: 0
US-09-397-967-16 (1-1099) x US-08-436-044-5 (1-3969)
QY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
Db 2298 GTCCACCGACGACGCTGCTCGCAACATCTAGTC 2333
RESULT 55
US-08-222-616-23
Sequence 23, Application US/08222616
Patent No. 5635177
GENERAL INFORMATION:
APPLICANT: Bennett, Brian D.
APPLICANT: Goeddel, David
APPLICANT: Lee, James M.
APPLICANT: Matthews, William
APPLICANT: Tsai, Siao Ping
APPLICANT: Wood, William I.
TITLE OF INVENTION: PROTEIN TYROSINE KINASE AGONIST
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/222,616
FILING DATE: 4-APR-1994
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/00586
FILING DATE: 22-JAN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/826935
FILING DATE: 22-JAN-1992
ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M.
REGISTRATION NUMBER:
REFERENCE/DOCKET NUMBER: 821P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1994
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 3969 bases
TYPE: nucleic acid

STRANDEDNESS: single
TOPOLOGY: linear
US-08-222-616-23

Alignment Scores:

Pred. No.:	0.338	Length:	3969
Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.09%	Indels:	0
DB:	1	Gaps:	0

US-09-397-967-16 (1-1099) x US-08-222-616-23 (1-3969)

OY 941 VALHISARGASPLEUALAALARGASNTLEUVAL 952
|||||

DB 2298 GTCCACCGAGACTGCTGCTGCAACATCTAGTC 2333

RESULT 56

US-08-436-054-5
Sequence 5, Application US/08436054

Patent No. 5864020

GENERAL INFORMATION:

APPLICANT: Bennett, Brian D.

APPLICANT: Matthews, William

TITLE OF INVENTION: HTR LIGAND

NUMBER OF SEQUENCES: 7

CORRESPONDENCE ADDRESS:

ADDRESSEE: Genentech, Inc.

STREET: 460 Point San Bruno Blvd

CITY: South San Francisco

STATE: California

COUNTRY: USA

ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: patin (Genentech)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/436.054

FILING DATE: 05-MAY-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/277722

FILING DATE: 20-JUL-1994

ATTORNEY/AGENT INFORMATION:

NAME: Lee, Wendy M.

REGISTRATION NUMBER: 00.000

REFERENCE/DOCKET NUMBER: 902D1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415/225-1994

TELEFAX: 415/952-9881

TELEX: 910/371-7168

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 3969 bases

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

US-08-436-054-5

Alignment Scores:

Pred. No.:	0.338	Length:	3969
Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.09%	Indels:	0
DB:	2	Gaps:	0

US-09-397-967-16 (1-1099) x US-08-436-054-5 (1-3969)

OY 941 VALHISARGASPLEUALAALARGASNTLEUVAL 952

DB 2298 GTCCACCGAGACTGCTGCTGCAACATCTAGTC 2333
|||||

RESULT 57

US-08-446-648-23
Sequence 23, Application US/08446648

Patent No. 631302

GENERAL INFORMATION:

APPLICANT: Genentech, Inc.

APPLICANT: Bennett, Brian D.

APPLICANT: Goeddel, David

APPLICANT: Lee, James M.

APPLICANT: Matthews, William

APPLICANT: Tsai, Shao Ping

APPLICANT: Wood, William I.

TITLE OF INVENTION: PROTEIN TYROSINE KINASE ACONIST ANTIBODIES

NUMBER OF SEQUENCES: 45

CORRESPONDENCE ADDRESS:

ADDRESSEE: Genentech, Inc.

STREET: 460 Point San Bruno Blvd

CITY: South San Francisco

STATE: California

COUNTRY: USA

ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: winpatin (Genentech)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/446.648

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/222616

FILING DATE: 04-APR-1994

ATTORNEY/AGENT INFORMATION:

NAME: Lee, Wendy M.

REGISTRATION NUMBER: 40,378

REFERENCE/DOCKET NUMBER: P0821P3PCT

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415/225-1994

TELEFAX: 415/952-9881

TELEX: 910/371-7168

INFORMATION FOR SEQ ID NO: 23:

SEQUENCE CHARACTERISTICS:

LENGTH: 3969 base pairs

TYPE: Nucleic Acid

STRANDEDNESS: Single

TOPOLOGY: linear

US-08-446-648-23

Alignment Scores:

Pred. No.:	0.338	Length:	3969
Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.09%	Indels:	0
DB:	4	Gaps:	0

US-09-397-967-16 (1-1099) x US-08-446-648-23 (1-3969)

OY 941 VALHISARGASPLEUALAALARGASNTLEUVAL 952

DB 2298 GTCCACCGAGACTGCTGCTGCAACATCTAGTC 2333

RESULT 58

PCT-US95-04228-23
Sequence 23, Application PC/TUS9504228

GENERAL INFORMATION:

APPLICANT: Genentech, Inc.

APPLICANT: Bennett, Brian D.

APPLICANT: Goeddel, David

APPLICANT: Lee, James M.
APPLICANT: Mathews, William
APPLICANT: Tsai, Siao Ping
APPLICANT: Wood, William I.
TITLE OF INVENTION: PROTEIN TYROSINE KINASE AGONIST ANTIBODIES
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/04228
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/222616
FILING DATE: 04-APR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Wendy M. Lee
REGISTRATION NUMBER: 00,000
REFERENCE/DOCKET NUMBER: 821P3PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1994
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 3969 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
PCT-US95-04228-23

Alignment Scores:
Pred. No.: 0.338 Length: 3969
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
Gaps: 0
DB: 5

US-09-397-967-16 (1-1099) x PCT-US95-04228-23 (1-3969)

OY 941 VALHISARGASPLEUAIAlAAlaRgAsnIleUeVal 952
DB 2298 GTCCACCGAGACCTGCTGCTGCAACATCTAGTC 2333

RESULT 59
PCT-US95-08812-5
Sequence 5, Application PC/TUS9508812
GENERAL INFORMATION:
APPLICANT: Genentech, Inc.
TITLE OF INVENTION: HTRK LIGAND
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/08812
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M.
REGISTRATION NUMBER: 00,000
REFERENCE/DOCKET NUMBER: 902PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1994
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 3969 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
PCT-US95-08812-5

Alignment Scores:
Pred. No.: 0.338 Length: 3969
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
Gaps: 0
DB: 5

US-09-397-967-16 (1-1099) x PCT-US95-08812-5 (1-3969)

OY 941 VALHISARGASPLEUAIAlAAlaRgAsnIleUeVal 952
DB 2298 GTCCACCGAGACCTGCTGCTGCAACATCTAGTC 2333

RESULT 60
US-08-348-143-3
Sequence 3, Application US/08348143
Patent No. 5506205
GENERAL INFORMATION:
APPLICANT: TAJIMA, HISAO
APPLICANT: KITAGAWA, KOICHIRO
APPLICANT: OHNO, HIROYUKI
APPLICANT: UENO, TOSHIO
TITLE OF INVENTION: A No. 5506205el Polypeptide of Protein p140 and DNAs
TITLE OF INVENTION: encoding it
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS
STREET: 2100 Pennsylvania Avenue, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20037-3202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/348,143
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 315806/1993
FILING DATE: 24-NOV-1993
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)293-7060
TELEFAX: (202)293-7860
TELEX: 6491103

;; INFORMATION FOR SEQ ID NO: 3:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 4027 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: cDNA to mRNA
;; ORIGINAL SOURCE:
;; ORGANISM: rat
;; TISSUE TYPE: skeletal muscle myoblast
;; CELL LINE: L6
US-08-348-143-3

Alignment Scores:
Pred. No.: 0.343 Length: 4027
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 1 Gaps: 0

US-09-397-967-16 (1-1099) x US-08-348-143-3 (1-4027)

OY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
Db 2509 GTGCACCGTGACCTCGCTGCCGCGACATCCTTGTC 2544

RESULT 61
US-08-348-143-4
; Sequence 4, Application US/08348143
; Patent No. 5506205
; GENERAL INFORMATION:
; APPLICANT: TAJIMA, HISAO
; APPLICANT: KITAGAWA, KOICHIRO
; APPLICANT: OHNO, HIROYUKI
; APPLICANT: UENO, TOSHIO
; TITLE OF INVENTION: A No. 5506205el Polypeptide of Protein p140 and DNAs
; NUMBER OF SEQUENCES: encoding 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS
; STREET: 2100 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20037-3202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/348,143
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 315806/1993
; FILING DATE: 24-NOV-1993
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)293-7060
; TELEFAX: (202)293-7860
; TELEX: 6491103
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4027 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; ORIGINAL SOURCE:
; ORGANISM: rat
; TISSUE TYPE: skeletal muscle myoblast
; CELL LINE: L6

;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: 262..3243
;; IDENTIFICATION METHOD: by similarity to some other pattern
US-08-348-143-4

Alignment Scores:
Pred. No.: 0.343 Length: 4027
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 1 Gaps: 0

US-09-397-967-16 (1-1099) x US-08-348-143-4 (1-4027)

OY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
Db 2509 GTGCACCGTGACCTCGCTGCCGCGACATCCTTGTC 2544

RESULT 62
US-08-571-785-3
; Sequence 3, Application US/08571785
; Patent No. 5804411
; GENERAL INFORMATION:
; APPLICANT: TAJIMA, HISAO
; APPLICANT: KITAGAWA, KOICHIRO
; APPLICANT: OHNO, HIROYUKI
; APPLICANT: UENO, TOSHIO
; TITLE OF INVENTION: A No. 5804411el Polypeptide of Protein p140
; NUMBER OF SEQUENCES: 16
; NUMBER OF SEQUENCES: and DNAs encoding 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS
; STREET: 2100 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20037-3202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/571,785
; FILING DATE: 13-DEC-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/348,143
; FILING DATE: 23-NOV-1994
; APPLICATION NUMBER: JP 315806/1993
; FILING DATE: 24-NOV-1993
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)293-7060
; TELEFAX: (202)293-7860
; TELEX: 6491103
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4027 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; ORIGINAL SOURCE:
; ORGANISM: rat
; TISSUE TYPE: skeletal muscle myoblast
; CELL LINE: L6
US-08-571-785-3

Alignment Scores:
Pred. No.: 0.343 Length: 4027
Score: 12.00 Matches: 12

Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 1.09%
DB: 1
Conservative: 0
Matches: 0
Indels: 0
Gaps: 0

US-09-397-967-16 (1-1099) x US-08-571-785-3 (1-4027)

Oy 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
Db 2509 GTGCACCGTGACCTCGCTGCCGCAACATCCTTGTG 2544

RESULT 63

US-08-571-785-4

Sequence 4, Application US/08571785
Patent No. 5804411

GENERAL INFORMATION:

APPLICANT: TAJIMA, HISAO

APPLICANT: KITAGAWA, KOICHIRO

APPLICANT: OHNO, HIROYUKI

APPLICANT: UENO, TOSHIO

TITLE OF INVENTION: A No. 5804411el Polypeptide of Protein p140

TITLE OF INVENTION: and DNAs encoding it

NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS:

ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS

STREET: 2100 Pennsylvania Avenue, N.W.

CITY: Washington

STATE: D.C.

COUNTRY: U.S.A.

ZIP: 20037-3202

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/571,785

FILING DATE: 13-DEC-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/348,143

FILING DATE: 23-NOV-1994

APPLICATION NUMBER: JP 315806/1993

FILING DATE: 24-NOV-1993

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202)293-7060

TELEFAX: (202)293-7860

TELEX: 6491103

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 4027 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA to mRNA

ORGANISM SOURCE:

ORGANISM: rat

TISSUE TYPE: skeletal muscle myoblast

CELL LINE: L6

FEATURE:

NAME/KEY: CDS

LOCATION: 262..3243

IDENTIFICATION METHOD: by similarity to some other pattern

Alignment Scores:

Pred. No.: 0.343
Score: 12.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 1.09%
DB: 1
Length: 4027
Matches: 12
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0

US-09-397-967-16 (1-1099) x US-08-571-785-4 (1-4027)

Oy 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
Db 2509 GTGCACCGTGACCTCGCTGCCGCAACATCCTTGTG 2544

RESULT 64

US-09-192-435-3

Sequence 3, Application US/09192435

Patent No. 6303320

GENERAL INFORMATION:

APPLICANT: TAJIMA, HISAO

APPLICANT: KITAGAWA, KOICHIRO

APPLICANT: OHNO, HIROYUKI

APPLICANT: UENO, TOSHIO

TITLE OF INVENTION: A No. 6303320el Polypeptide of Protein p140

TITLE OF INVENTION: and DNAs encoding it

NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS:

ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS

STREET: 2100 Pennsylvania Avenue, N.W.

CITY: Washington

STATE: D.C.

COUNTRY: U.S.A.

ZIP: 20037-3202

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/192,435

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/571,785

FILING DATE: 13-DEC-1995

APPLICATION NUMBER: 08/348,143

FILING DATE: 23-NOV-1994

APPLICATION NUMBER: JP 315806/1993

FILING DATE: 24-NOV-1993

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202)293-7060

TELEFAX: (202)293-7860

TELEX: 6491103

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 4027 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA to mRNA

ORGANISM SOURCE:

ORGANISM: rat

TISSUE TYPE: skeletal muscle myoblast

CELL LINE: L6

FEATURE:

NAME/KEY: CDS

LOCATION: 262..3243

IDENTIFICATION METHOD: by similarity to some other pattern

Alignment Scores:

Pred. No.: 0.343
Score: 12.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 1.09%
DB: 4
Length: 4027
Matches: 12
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0

US-09-397-967-16 (1-1099) x US-09-192-435-3 (1-4027)

Oy 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
Db 2509 GTGCACCGTGACCTCGCTGCCGCAACATCCTTGTG 2544

RESULT 65

US-09-192-435-4
; Sequence 4, Application US/09192435
; Patent No. 6303320
GENERAL INFORMATION:
APPLICANT: TAJIMA, HISAO
APPLICANT: KITAGAWA, KOICHIRO
APPLICANT: OHNO, HIROYUKI
TITLE OF INVENTION: A No. 6303320el Polypeptide of Protein p140
TITLE OF INVENTION: and DNAs encoding it
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS
STREET: 2100 Pennsylvania Avenue, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20037-3202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/192,435
FILING DATE:
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/571,785
FILING DATE: 13-DEC-1995
APPLICATION NUMBER: 08/348,143
FILING DATE: 23-NOV-1994
APPLICATION NUMBER: JP 315806/1993
FILING DATE: 24-NOV-1993
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)293-7060
TELEFAX: (202)293-7860
TELEX: 6491103
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 4027 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
ORIGINAL SOURCE:
ORGANISM: rat
TISSUE TYPE: skeletal muscle myoblast
CELL LINE: L6
FEATURE:
NAME/KEY: CDS
LOCATION: 262..3243
IDENTIFICATION METHOD: by similarity to some other pattern
US-09-192-435-4
Alignment Scores:
Pred. No.: 0.343 Length: 4027
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 4 Gaps: 0
US-09-397-967-16 (1-1099) x US-09-192-435-4 (1-4027)
Oy 941 ValHisArgAspLeuAlaIaIaArgAsnIleLeuVal 952
|||||
Db 2509 GTGACACGTGACCTCGCTCCGCAACATCTTGTC 2544
RESULT 66
US-09-558-340-3
; Sequence 3, Application US/09558340
; Patent No. 6432913

GENERAL INFORMATION:
APPLICANT: TAJIMA, HISAO
APPLICANT: KITAGAWA, KOICHIRO
APPLICANT: OHNO, HIROYUKI
TITLE OF INVENTION: A No. 6432913el Polypeptide of Protein p140
TITLE OF INVENTION: and DNAs encoding it
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS
STREET: 2100 Pennsylvania Avenue, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20037-3202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/558,340
FILING DATE: 26-APR-2000
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 09/192,435
FILING DATE: 08-JAN-1998
APPLICATION NUMBER: 08/571,785
FILING DATE: 13-DEC-1995
APPLICATION NUMBER: 08/348,143
FILING DATE: 23-NOV-1994
APPLICATION NUMBER: JP 315806/1993
FILING DATE: 24-NOV-1993
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)293-7060
TELEFAX: (202)293-7860
TELEX: 6491103
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 4027 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
ORIGINAL SOURCE:
ORGANISM: rat
TISSUE TYPE: skeletal muscle myoblast
CELL LINE: L6
US-09-558-340-3
Alignment Scores:
Pred. No.: 0.343 Length: 4027
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 4 Gaps: 0
US-09-397-967-16 (1-1099) x US-09-558-340-3 (1-4027)
Oy 941 ValHisArgAspLeuAlaIaIaArgAsnIleLeuVal 952
|||||
Db 2509 GTGACACGTGACCTCGCTCCGCAACATCTTGTC 2544
RESULT 67
US-09-558-340-4
; Sequence 4, Application US/09558340
; Patent No. 6432913
GENERAL INFORMATION:
APPLICANT: TAJIMA, HISAO
APPLICANT: KITAGAWA, KOICHIRO
APPLICANT: OHNO, HIROYUKI
TITLE OF INVENTION: A No. 6432913el Polypeptide of Protein p140
TITLE OF INVENTION: and DNAs encoding it
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS
STREET: 2100 Pennsylvania Avenue, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20037-3202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/558,340
FILING DATE: 26-APR-2000
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 09/192,435
FILING DATE: 08-JAN-1998
APPLICATION NUMBER: 08/571,785
FILING DATE: 13-DEC-1995
APPLICATION NUMBER: 08/348,143
FILING DATE: 23-NOV-1994
APPLICATION NUMBER: JP 315806/1993
FILING DATE: 24-NOV-1993
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)293-7060
TELEFAX: (202)293-7860
TELEX: 6491103
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 4027 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
ORIGINAL SOURCE:
ORGANISM: rat
TISSUE TYPE: skeletal muscle myoblast
CELL LINE: L6
US-09-558-340-3

;; TITLE OF INVENTION: A No. 6432913el Polypeptide of Protein p140
;; TITLE OF INVENTION: and DNAs encoding it
;; NUMBER OF SEQUENCES: 16
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: SUGHRUE, MIOW, ZINN, MACPHEAK & SEAS
;; STREET: 2100 Pennsylvania Avenue, N.W.
;; CITY: Washington
;; STATE: D.C.
;; COUNTRY: U.S.A.
;; ZIP: 20037-3202
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/558,340
;; FILING DATE: 26-APR-2000
;; CLASSIFICATION:
;; PRIORITY APPLICATION DATA:
;; APPLICATION NUMBER: 09/192,435
;; FILING DATE: 08-JAN-1998
;; TELECOMMUNICATION INFORMATION:
;; APPLICATION NUMBER: 08/571,785
;; FILING DATE: 13-DEC-1995
;; APPLICATION NUMBER: 08/348,143
;; FILING DATE: 23-NOV-1994
;; APPLICATION NUMBER: JP 315806/1993
;; FILING DATE: 24-NOV-1993
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (202)293-7060
;; TELEFAX: (202)293-7860
;; TELEX: 6491103
;; INFORMATION FOR SEQ. ID NO: 4:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 4027 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: cDNA to mRNA
;; ORIGINAL SOURCE:
;; ORGANISM: rat
;; TISSUE TYPE: skeletal muscle myoblast
;; CELL LINE: L6
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: 262..3243
;; IDENTIFICATION METHOD: by similarity to some other pattern
US-09-558-340-4

Alignment Scores:
Pred. No.: 0.343 Length: 4027
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 4 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-558-340-4 (1-4027)
QY 941 VALHISARGASPLEUALAIAARGANILEUVAL 952
DB 2509 GTGCACCGTGACCTGCTGCCGACACATCTGTTC 2344

RESULT 68
US-08-162-809-17
; Sequence 17, Application US/08162809
; Patent No. 5457048
; GENERAL INFORMATION:
; APPLICANT: Pasquale, Elena B.
; APPLICANT: Sajjadi, Fereydoun G.
; TITLE OF INVENTION: NOVEL EPH-RELATED TYROSINE KINASES,
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES, AND METHODS OF USE
; NUMBER OF SEQUENCES: 26

;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: CAMPBELL AND FLORES
;; STREET: 4370 La Jolla Village Drive, Suite 700
;; CITY: San Diego
;; STATE: California
;; COUNTRY: United States of America
;; ZIP: 92122
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/162,809
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Campbell, Cathryn A.
;; REGISTRATION NUMBER: P-LJ 9503
;; REFERENCE/DOCKET NUMBER: 31,815
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (619) 535-9001
;; TELEFAX: (619) 535-8949
;; INFORMATION FOR SEQ. ID NO: 17:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 4049 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: both
;; TOPOLOGY: linear
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: 10..2994
US-08-162-809-17

Alignment Scores:
Pred. No.: 0.344 Length: 4049
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 1 Gaps: 0

US-09-397-967-16 (1-1099) x US-08-162-809-17 (1-4049)
QY 941 VALHISARGASPLEUALAIAARGANILEUVAL 952
DB 2263 GTGCACCGGACCTGCTGCCGACACATCTGTTC 2298

RESULT 69
US-08-162-809-11
; Sequence 11, Application US/08162809
; Patent No. 5457048
; GENERAL INFORMATION:
; APPLICANT: Pasquale, Elena B.
; APPLICANT: Sajjadi, Fereydoun G.
; TITLE OF INVENTION: NOVEL EPH-RELATED TYROSINE KINASES,
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES, AND METHODS OF USE
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CAMPBELL AND FLORES
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States of America
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/162,809
; FILING DATE:

CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LJ 9503
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 4097 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 10..3042
US-08-162-809-11

Alignment Scores:
Pred. No.: 0.348 Length: 4097
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
Gaps: 0
DB:

US-09-397-967-16 (1-1099) x US-08-162-809-11 (1-4097)

OY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
Db 2311 GTGACCGGAGACTGCTGCCGCAACATCCTGTC 2346

RESULT 70
US-08-368-776A-1
Sequence 1, Application US/08368776A
Patent No. 6300482
GENERAL INFORMATION:
APPLICANT: Ciossek, Thomas
APPLICANT: Ullrich, Axel
TITLE OF INVENTION: METHODS FOR DIAGNOSIS
TITLE OF INVENTION: AND TREATMENT OF MDK1
TITLE OF INVENTION: SIGNAL TRANSDUCTION
TITLE OF INVENTION: DISORDERS
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 MB
MEDIUM TYPE: storage
COMPUTER: IBM compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/368,776A
FILING DATE: January 3, 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
PRIOR APPLICATION DATA: including application
PRIOR APPLICATION DATA: described below:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 208/007

none

TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4304 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: nucleic
US-08-368-776A-1

Alignment Scores:
Pred. No.: 0.364 Length: 4304
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
Gaps: 0
DB:

US-09-397-967-16 (1-1099) x US-08-368-776A-1 (1-4304)

OY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
Db 2495 GTTCACAGGAGCCTTGCGCGCGCAACATCCTGTC 2530

RESULT 71
PCT-US96-00419-1
Sequence 1, Application PC/TUS9600419
GENERAL INFORMATION:
APPLICANT: Thomas Ciossek, Axel Ullrich, Birgit
APPLICANT: Millaueer
TITLE OF INVENTION: METHODS FOR DIAGNOSIS AND TREATMENT OF
TITLE OF INVENTION: MDK1 SIGNAL TRANSDUCTION DISORDERS
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 MB
COMPUTER: IBM compatible
OPERATING SYSTEM: IBM P.C. DOS (Version 5.0)
SOFTWARE: WordPerfect (Version 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/00419
FILING DATE: January 3, 1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
PRIOR APPLICATION DATA: including application
PRIOR APPLICATION DATA: described below:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 208/007
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4304
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: nucleic
PCT-US96-00419-1

none

```
Alignment Scores:
Pred. No.: 0.364 Length: 4304
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 5 Gaps: 0

US-09-397-967-16 (1-1099) x PCT-US96-00419-1 (1-4304)
OY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
Db 2495 GTTCACAGGACCTTGCGAGCGCGCAACATCTCTGTC 2530

RESULT 72
PCT-US93-06251-34
Sequence 34, Application PC/TUS9306251
GENERAL INFORMATION:
APPLICANT: Wicksrom, Eric and Rife, Jason P.
TITLE OF INVENTION: Trivalent Synthesis of Oligonucleotides Containing
NUMBER OF SEQUENCES: 93
CORRESPONDENCE ADDRESS:
ADDRESS: SCULLY, SCOTT, MURPHY & PRESSER
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: NY
COUNTRY: USA
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/06251
FILING DATE: 19930630
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: DIGILIO, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 8586
TELECOMMUNICATION INFORMATION:
TELEPHONE: 516-742-4343
TELEFAX: 516-742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 4508 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
PCT-US93-06251-34

Alignment Scores:
Pred. No.: 0.38 Length: 4508
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 5 Gaps: 0

US-09-397-967-16 (1-1099) x PCT-US93-06251-34 (1-4508)
OY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
Db 2542 GTTCATCGGACCTTGCGAGCGCGCAACATCTCTGTA 2577

RESULT 73
US-08-449-645A-16
Sequence 16, Application US/08449645A

Alignment Scores:
Pred. No.: 0.381 Length: 4529
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 2 Gaps: 0

US-09-397-967-16 (1-1099) x US-08-449-645A-16 (1-4529)
OY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
Db 2448 GTTCACAGGACCTTGCGAGCTCGCAATATCTCTGTC 2483

RESULT 74
US-08-702-367A-16
Sequence 16, Application US/08702367A
Patent No. 5981246
GENERAL INFORMATION:
APPLICANT: Fox, Gary M.
TITLE OF INVENTION: Eph-Like Receptor Protein Tyrosine
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESS: Amgen Patent Operations/RBW
STREET: 1840 Dehavilland Drive
CITY: Thousand Oaks
STATE: California
COUNTRY: USA
ZIP: 91320
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
```

APPLICATION NUMBER: US/08/702,367A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Winter, Robert B.
REFERENCE/DOCKET NUMBER: A-287
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 4529 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 186..3182
US-08-702-367A-16

Alignment Scores:
Pred. No.: 0.381 Length: 4529
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
Gaps: 0
DB:

US-09-397-967-16 (1-1099) x US-08-702-367A-16 (1-4529)
OY 941 VALHISARGASPLEUALAALARGANILLEUVAL 952
DB 2448 GTTCACAGGACCTTGACGCTGCATATTTCTGTGC 2483

RESULT 75
PCT-US95-04681-16
Sequence 16, Application PC/TUS9504681
GENERAL INFORMATION:
APPLICANT: Fox, Gary M.
TITLE OF INVENTION: Eph-Like Receptor Protein Tyrosine
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amgen Patent Operations/RMW
STREET: 1840 Dehaviiland Drive
CITY: Thousand Oaks
STATE: California
COUNTRY: USA
ZIP: 91320
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/04681
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Winter, Robert B.
REFERENCE/DOCKET NUMBER: A-287
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 4529 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 186..3182
PCT-US95-04681-16

Alignment Scores:
Pred. No.: 0.381 Length: 4529

Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
Gaps: 0
DB:

US-09-397-967-16 (1-1099) x PCT-US95-04681-16 (1-4529)
OY 941 VALHISARGASPLEUALAALARGANILLEUVAL 952
DB 2448 GTTCACAGGACCTTGACGCTGCATATTTCTGTGC 2483

RESULT 76
US-08-857-076-11
Sequence 11, Application US/08857076C
GENERAL INFORMATION:
APPLICANT: Ruvkun, Gary
APPLICANT: Kimura, Koutarou
APPLICANT: Patterson, Garth
APPLICANT: Ogil, Scott
APPLICANT: Tissenbaum, Heidi
APPLICANT: Morris, Jason
APPLICANT: Kowsek, Allison
TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC TOOLS FOR
FILE REFERENCE: 00786/351001
CURRENT FILING DATE: 1997-05-15
NUMBER OF SEQ ID NOS: 114
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11
LENGTH: 5816
TYPE: DNA
ORGANISM: Caenorhabditis elegans
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)..(5816)
OTHER INFORMATION: n = A,T,C or G
US-08-857-076-11

Alignment Scores:
Pred. No.: 0.479 Length: 5816
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
Gaps: 0
DB:

US-09-397-967-16 (1-1099) x US-08-857-076-11 (1-5816)
OY 1003 SeraspVALTTPSERPHGGLYVALLEUVAL 1014
DB 4455 TGTGATGTTGGAGCTTGCGAGTGTCTCTATGA 4490

RESULT 77
US-09-741-154-3
Sequence 3, Application US/09741154
GENERAL INFORMATION:
APPLICANT: BEASLEY, Ellen M. et al
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
FILE REFERENCE: C1001061
CURRENT FILING DATE: 2000-12-21
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 16389
TYPE: DNA

ORGANISM: Human
US-09-741-154-3

Alignment Scores:
Pred. No.: 1.23 Length: 16389
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 4 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-741-154-3 (1-16389)

Qy 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
Db 13220 GTGCACCCGACGCTGCGCCGCGCAACATCCTGCTC 13255

RESULT 78
US-08-222-616-7

Sequence 7, Application US/08222616
Patent No. 5635177
GENERAL INFORMATION:
APPLICANT: Bennett, Brian D.
APPLICANT: Goeddel, David
APPLICANT: Lee, James M.
APPLICANT: Matthews, William
APPLICANT: Tsai, Siao Ping
APPLICANT: Wood, William I.
TITLE OF INVENTION: PROTEIN TYROSINE KINASE AGONIST
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080

COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 kb floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: paln (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/222,616
FILING DATE: 4-APR-1994
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/00586
FILING DATE: 22-JAN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/826935
FILING DATE: 22-JAN-1992
ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M.
REGISTRATION NUMBER:
REFERENCE/DOCKET NUMBER: 821P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1994
TELEFAX: 415/952-9881
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 147 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-222-616-7

Alignment Scores:
Pred. No.: 0.137 Length: 147
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.00% Indels: 0
DB: 1 Gaps: 0

US-09-397-967-16 (1-1099) x US-08-222-616-7 (1-147)

Qy 942 HisArgAspLeuAlaAlaArgAsnIleLeuVal 952
Db 10 CACAGAGCCTAGCAGCAGCGCAACATCCTGCTC 42

RESULT 79
US-08-876-882-3

Sequence 3, Application US/08876882
Patent No. 5981201
GENERAL INFORMATION:
APPLICANT: Avraham, Hava
APPLICANT: Groopman, Jerome E.
TITLE OF INVENTION: METHODS OF DETECTION AND TREATMENT
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds P.C.
STREET: Two Mallitia Drive
CITY: Lexington
STATE: MA
COUNTRY: USA
ZIP: 02173-4799

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows
SOFTWARE: FASTSEQ for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/876,882
FILING DATE: 16-JUN-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/035,228
FILING DATE: 08-JAN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Doreen, Hoyle M.
REGISTRATION NUMBER: 36,361
REFERENCE/DOCKET NUMBER: NEDH97-01PA
TELECOMMUNICATION INFORMATION:
TELEPHONE: 781-861-6240
TELEFAX: 781-861-9540
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 147 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-876-882-3

Alignment Scores:
Pred. No.: 0.137 Length: 147
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.00% Indels: 0
DB: 2 Gaps: 0

US-09-397-967-16 (1-1099) x US-08-876-882-3 (1-147)

Qy 942 HisArgAspLeuAlaAlaArgAsnIleLeuVal 952
Db 10 CACAGAGCCTAGCAGCAGCGCAACATCCTGCTC 42

RESULT 80
US-08-446-648-7

Sequence 7, Application US/08446648
Patent No. 6331302

GENERAL INFORMATION:
APPLICANT: Genentech, Inc.
APPLICANT: Bennett, Brian D.
APPLICANT: Goeddel, David
APPLICANT: Lee, James M.
APPLICANT: Matthews, William
APPLICANT: Tsai, Siao Ping
APPLICANT: Wood, William I.
TITLE OF INVENTION: PROTEIN TYROSINE KINASE AGONIST ANTIBODIES
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 MB floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Winpatin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/446,648
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/222616
FILING DATE: 04-APR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M.
REGISTRATION NUMBER: 40,378
REFERENCE/DOCKET NUMBER: P0821P3PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1994
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 147 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Single
TOPOLOGY: Linear
US-08-446-648-7

Alignment Scores:
Pred. No.: 0.137 Length: 147
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.00% Indels: 0
DB: 4 Gaps: 0

US-09-397-967-16 (1-1099) x US-08-446-648-7 (1-147)
QY 942 HisArgAspLeuAlaAlaArgAsnIleuVal 952
DB 10 CACAGAGACTAGCAGCAGCAGCAGCATCTGTC 42

RESULT 81
US-09-315-928-3
Sequence 3, Application US/09315928
Patent No. 6368796
GENERAL INFORMATION:
APPLICANT: Avraham, Hava
APPLICANT: Groopman, Jerome E.
TITLE OF INVENTION: METHODS OF DETECTION AND TREATMENT OF
FILE REFERENCE: NED97-01P4Z
CURRENT APPLICATION NUMBER: US/09/315,928
CURRENT FILING DATE: 1999-05-20
PRIOR APPLICATION NUMBER: US 08/876,882
PRIOR FILING DATE: 1997-06-16

PRIOR APPLICATION NUMBER: US 60/035,228
PRIOR FILING DATE: 1997-01-08
NUMBER OF SEQ ID NOS: 5
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 3
LENGTH: 147
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (1)...(150)
US-09-315-928-3

Alignment Scores:
Pred. No.: 0.137 Length: 147
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.00% Indels: 0
DB: 4 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-315-928-3 (1-147)
QY 942 HisArgAspLeuAlaAlaArgAsnIleuVal 952
DB 10 CACAGAGACTAGCAGCAGCAGCAGCATCTGTC 42

RESULT 82
PCT-US95-04228-7
Sequence 7, Application PC/TUS9504228
GENERAL INFORMATION:
APPLICANT: Genentech, Inc.
APPLICANT: Bennett, Brian D.
APPLICANT: Goeddel, David
APPLICANT: Lee, James M.
APPLICANT: Matthews, William
APPLICANT: Tsai, Siao Ping
APPLICANT: Wood, William I.
TITLE OF INVENTION: PROTEIN TYROSINE KINASE AGONIST ANTIBODIES
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 KB floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/04228
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/222616
FILING DATE: 04-APR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Wendy M. Lee
REGISTRATION NUMBER: 00,000
REFERENCE/DOCKET NUMBER: 821P3PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1994
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 147 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

PCT-US95-04228-7

Alignment Scores:

Pred. No.:	0.137	Length:	147
Score:	11.00	Matches:	11
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.00%	Indels:	0
DB:	5	Gaps:	0

US-09-397-967-16 (1-1099) x PCT-US95-04228-7 (1-147)

QY 942 HisargaspLeuAlaalaArgAsnIleuVal 952

Db 10 CACAGAGACCTAGCAGCAGCAGCATCTGTC 42

RESULT 83

US-08-278-089A-3

Sequence 3, Application US/08278089A

Patent No. 5681714

GENERAL INFORMATION:

APPLICANT: Breilman, Martin L.

APPLICANT: Rosant, Janet

APPLICANT: Dumont, Daniel J.

APPLICANT: Yamaguchi, Terry P.

TITLE OF INVENTION: No. 5681714el Receptor Tyrosine Kinase

NUMBER OF SEQUENCES: 33

CORRESPONDENCE ADDRESS:

ADDRESSEE: Bereskin & Parr

STREET: 40 King Street West

CITY: Toronto

STATE: Ontario

COUNTRY: Canada

ZIP: M5H 3Y2

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/278,089A

FILING DATE: 20-JUL-1994

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: Kurdyak, Linda M.

REGISTRATION NUMBER: 34,971

REFERENCE/DOCKET NUMBER: 3153-111

TELECOMMUNICATION INFORMATION:

TELEPHONE: (416) 364-7311

TELEFAX: (416) 361-1398

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 1590 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

ORIGINAL SOURCE:

ORGANISM: Mus musculus

DEVELOPMENTAL STAGE: Embryo

IMMEDIATE SOURCE:

LIBRARY: murine embryonic lambda gt10 cDNA library

CLONE: 1.6kb clone

POSITION IN GENOME:

CHROMOSOME/SEGMENT: 4

MAP POSITION: Between the brown and pmv-23 loci

FEATURE:

NAME/KEY: CDS

LOCATION: 1..903

US-08-278-089A-3

Alignment Scores:

Pred. No.: 1.19

Length: 1590

Score: 11.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 1.00%

DB: 1

US-09-397-967-16 (1-1099) x US-08-278-089A-3 (1-1590)

QY 942 HisargaspLeuAlaalaArgAsnIleuVal 952

Db 415 CACAGGACCTGGCTGCCAGAACATTTAGTT 447

RESULT 84

US-07-934-393B-3

Sequence 3, Application US/07934393B

Patent No. 546596

GENERAL INFORMATION:

APPLICANT: BREITMAN, MARTIN L.

APPLICANT: DUMONT, DANIEL

APPLICANT: GRADWOHL, GERARD G.

TITLE OF INVENTION: TISSUE SPECIFIC TRANSCRIPTIONAL

NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESS:

ADDRESSEE: BERESKIN & PARR

STREET: 40 King Street West

CITY: Toronto

STATE: Ontario

COUNTRY: Canada

ZIP: M5H 3Y2

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/934,393B

FILING DATE: 25-AUG-1992

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Kurdyak, Linda M.

REGISTRATION NUMBER: 34,971

REFERENCE/DOCKET NUMBER: 3153-64

TELECOMMUNICATION INFORMATION:

TELEPHONE: (416) 354-7311

TELEFAX: (416) 361-1398

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 1601 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

ORIGINAL SOURCE:

ORGANISM: Mus pahari

DEVELOPMENTAL STAGE: Embryo

IMMEDIATE SOURCE:

LIBRARY: murine embryonic lambda gt10 cDNA library

CLONE: 1.6kb clone

POSITION IN GENOME:

CHROMOSOME/SEGMENT: 4

MAP POSITION: Between the brown and pmv-23 loci

FEATURE:

NAME/KEY: CDS

LOCATION: 1..903

US-07-934-393B-3

Alignment Scores:

Pred. No.: 1.2

Length: 1601

Score: 11.00

Matches: 11

Percent Similarity: 100.00%

Conservative: 0

Best Local Similarity: 100.00%

Mismatches: 0

Query Match: 1.00%

Indels: 0

ADDRESSEE: CAMPBELL AND FLORES
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: United States of America
ZIP: 92122

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/162,809
FILING DATE:
CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LJ 9503
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3059 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear

FEATURE:
NAME/KEY: CDS
LOCATION: 2..2167
US-08-162-809-3

Alignment Scores:
Pred. No.: 2.16 Length: 3059
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.00% Indels: 0
DB: 1 Gaps: 0

US-09-397-967-16 (1-1099) x US-08-162-809-3 (1-3059)

QY 941 VALHISARGASPLEUALAALARGASNIIELEU 951
|||||

Db 1445 GTACACAGACCTCGCTGCCAGATATCTCTC 1477

RESULT 88
US-08-162-809-19
Sequence 19, Application US/08162809
Patent No. 5457048
GENERAL INFORMATION:
APPLICANT: Pasquale, Elena B.
APPLICANT: Sajjad, Pateydon G.
TITLE OF INVENTION: NOVEL EPH-RELATED TYROSINE KINASES,
TITLE OF INVENTION: NUCLEOTIDE SEQUENCES, AND METHODS OF USE
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: CAMPBELL AND FLORES
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: United States of America
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/162,809
FILING DATE:
CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LJ 9503
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 3125 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear

FEATURE:
NAME/KEY: CDS
LOCATION: 2..2233
US-08-162-809-19

Alignment Scores:
Pred. No.: 2.2 Length: 3125
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.00% Indels: 0
DB: 1 Gaps: 0

US-09-397-967-16 (1-1099) x US-08-162-809-19 (1-3125)

QY 941 VALHISARGASPLEUALAALARGASNIIELEU 951
|||||

Db 1511 GTACACAGACCTCGCTGCCAGATATCTCTC 1543

RESULT 89
US-08-167-919A-9
Sequence 9, Application US/08167919A
Patent No. 5674691
GENERAL INFORMATION:
APPLICANT: Boyd, Andrew W.
APPLICANT: Simpson, Richard J.
APPLICANT: Wicks, Ian
APPLICANT: Ward, Larry D.
APPLICANT: Wilkinson, David
TITLE OF INVENTION: A NOVEL RECEPTOR-TYPE TYROSINE KINASE
TITLE OF INVENTION: AND USE THEREOF
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scully, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: U.S.A.
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/167,919A
FILING DATE: 18-APR-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PK6841 (AU)
FILING DATE: 21-JUN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PK9992 (AU)
FILING DATE: 12-DEC-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/AU92/00294
FILING DATE: 19-JUN-1992
ATTORNEY/AGENT INFORMATION:
NAME: Digilio, Frank S.
REGISTRATION NUMBER: 31,346

REFERENCE/DOCKET NUMBER: 9159
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 3132 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 100..3048
US-08-167-919A-9

Alignment Scores:
Pred. No.: 2.21 Length: 3132
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.00% Indels: 0
DB: 1 Gaps: 0

US-09-397-967-16 (1-1099) x US-08-167-919A-9 (1-3132)

OY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeu 951
DB 2326 GTTCACCGAGACCTGCTCGAACAATCTTG 2358

RESULT 90
US-08-715-106-9
Sequence 9, Application US/08715106
Patent No. 6020306
GENERAL INFORMATION:
APPLICANT: Boyd, Andrew W.
APPLICANT: Simpson, Richard J.
APPLICANT: Wicks, Ian
APPLICANT: Ward, Larry D.
APPLICANT: Wilkinson, David
TITLE OF INVENTION: A NOVEL RECEPTOR-TYPE TYROSINE KINASE
TITLE OF INVENTION: AND USE THEREOF
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scully, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: U.S.A.
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/715,106
FILING DATE: 18-SEP-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/167,919
FILING DATE: 18-APR-1994
APPLICATION NUMBER: PK6841 (AU)
FILING DATE: 21-JUN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PK9992 (AU)
FILING DATE: 12-DEC-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/AU92/00294
FILING DATE: 19-JUN-1992
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.

REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 9159
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 3132 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 100..3048
US-08-715-106-9

Alignment Scores:
Pred. No.: 2.21 Length: 3132
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.00% Indels: 0
DB: 3 Gaps: 0

US-09-397-967-16 (1-1099) x US-08-715-106-9 (1-3132)

OY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeu 951
DB 2326 GTTCACCGAGACCTGCTCGAACAATCTTG 2358

RESULT 91
US-08-449-645A-12
Sequence 12, Application US/08449645A
Patent No. 5981245
GENERAL INFORMATION:
APPLICANT: Fox, Gary M.
TITLE OF INVENTION: EPH-Like Receptor Protein Tyrosine
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amgen Patent Operations/RBW
STREET: 1840 Dehavenland Drive
CITY: Thousand Oaks
STATE: California
COUNTRY: USA
ZIP: 91320
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/449,645A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Winter, Robert B.
REFERENCE/DOCKET NUMBER: A-287
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 3162 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..2976
US-08-449-645A-12
Alignment Scores:

Pred. No.: 2.23 Length: 3162
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.00% Indels: 0
DB: 2 Gaps: 0

US-09-397-967-16 (1-1099) x US-08-449-645A-12 (1-3162)

Qy 941 ValHisArgAspLeuAlaAlaArgAsnIleLeu 951
|||||
Db 2251 GTGCATAGAGATCTTGCTGCCAGAAACATCTTA 2283

RESULT 92
US-08-702-367A-12
Sequence 12, Application US/08702367A
Patent No. 5981246
GENERAL INFORMATION:
APPLICANT: FOX, Gary M.
TITLE OF INVENTION: EPH-Like Receptor Protein Tyrosine
NUMBER OF INVENTION: Kinases
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amgen Patent Operations/RBM
STREET: 1840 Dehavilland Drive
CITY: Thousand Oaks
STATE: California
COUNTRY: USA
ZIP: 91320
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/702.367A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Winter, Robert B.
REFERENCE/DOCKET NUMBER: A-287
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 3162 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..2976
US-08-702-367A-12

Alignment Scores:
Pred. No.: 2.23 Length: 3162
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.00% Indels: 0
DB: 2 Gaps: 0

US-09-397-967-16 (1-1099) x US-08-702-367A-12 (1-3162)

Qy 941 ValHisArgAspLeuAlaAlaArgAsnIleLeu 951
|||||
Db 2251 GTGCATAGAGATCTTGCTGCCAGAAACATCTTA 2283

RESULT 93
PCT-US95-04681-12
Sequence 12, Application PC/7US9504681
GENERAL INFORMATION:
APPLICANT: Fox, Gary M.
TITLE OF INVENTION: EPH-Like Receptor Protein Tyrosine

TITLE OF INVENTION: Kinases
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amgen Patent Operations/RBM
STREET: 1840 Dehavilland Drive
CITY: Thousand Oaks
STATE: California
COUNTRY: USA
ZIP: 91320
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/04681.
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Winter, Robert B.
REFERENCE/DOCKET NUMBER: A-287
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 3162 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..2976
PCT-US95-04681-12

Alignment Scores:
Pred. No.: 2.23 Length: 3162
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.00% Indels: 0
DB: 5 Gaps: 0

US-09-397-967-16 (1-1099) x PCT-US95-04681-12 (1-3162)

Qy 941 ValHisArgAspLeuAlaAlaArgAsnIleLeu 951
|||||
Db 2251 GTGCATAGAGATCTTGCTGCCAGAAACATCTTA 2283

RESULT 94
US-08-162-809-15
Sequence 15, Application US/08162809
Patent No. 5457048
GENERAL INFORMATION:
APPLICANT: Pasquale, Elena B.
TITLE OF INVENTION: NOVEL EPH-RELATED TYROSINE KINASES,
NUMBER OF INVENTION: NUCLEOTIDE SEQUENCES, AND METHODS OF USE
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: CAMPBELL AND FLORES
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: United States of America
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/162,809
FILING DATE:
CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LJ 9503
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9901
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 3254 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 32..2980
US-08-162-809-15

Alignment Scores:
Pred. No.: 2.29 Length: 3254
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.00% Indels: 0
DB: 1 Gaps: 0

US-09-397-967-16 (1-1099) x US-08-162-809-15 (1-3254)

OY 941 VALHISARGASPLEUAIAlaArgAsnIleLeu 951

Db 2258 GTCCACGAGATCTGCTGCTGTAATATACTC 2290

RESULT 95

US-08-469-537A-102
Sequence 102, Application US/08469537A

Patent No. 5843749

GENERAL INFORMATION:

APPLICANT: Maisongier, et al.

TITLE OF INVENTION: EHK AND FOR TYROSINE

TITLE OF INVENTION: KINASES

NUMBER OF SEQUENCES: 107

CORRESPONDENCE ADDRESS:

ADDRESS: Regeneron Pharmaceuticals, Inc.

STREET: 777 Old Saw Mill River Road

CITY: Tarrytown

STATE: NY

COUNTRY: U.S.A.

ZIP: 10591

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/469,537A

FILING DATE: 06-JUN-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: USSN 08/406,247

FILING DATE: 17-MAR-1995

APPLICATION NUMBER: USSN 08/144,992

FILING DATE: 28-OCT-1993

APPLICATION NUMBER: USSN 07/736,559

FILING DATE: 26-JUL-1991

ATTORNEY/AGENT INFORMATION:

NAME: Kempster, Ph.D., Gall M

REGISTRATION/DOCKET NUMBER: 32,143

TELECOMMUNICATION INFORMATION:

TELEPHONE: 914-345-7400

TELEFAX: 914-345-7721

TELEX:
INFORMATION FOR SEQ ID NO: 102:

SEQUENCE CHARACTERISTICS:
LENGTH: 3906 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA
FEATURE:
NAME/KEY: CDS
LOCATION: 476..3493
US-08-469-537A-102

Alignment Scores:
Pred. No.: 2.7 Length: 3906
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.00% Indels: 0
DB: 2 Gaps: 0

US-09-397-967-16 (1-1099) x US-08-469-537A-102 (1-3906)

OY 941 VALHISARGASPLEUAIAlaArgAsnIleLeu 951

Db 2870 GTCCACGAGAGCTTGTCTGCTAGAACATCTTA 2902

RESULT 96

US-08-323-474-1
Sequence 1, Application US/08323474

Patent No. 5447860

GENERAL INFORMATION:

APPLICANT: Ziegler, Steven F.

TITLE OF INVENTION: NOVEL TYROSINE KINASE

NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESS:

ADDRESSEE: Immunex Corporation

STREET: 51 University Street

CITY: Seattle

STATE: Washington

COUNTRY: US

ZIP: 98101

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/323,474

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/905,600

FILING DATE: 26-JUN-1992

ATTORNEY/AGENT INFORMATION:

NAME: Seese, Cathryn A.

REGISTRATION NUMBER: 32,172

REFERENCE/DOCKET NUMBER: 2609

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 587-0430

TELEFAX: (206) 233-0644

TELEX: 756822

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 4138 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA to mRNA

FEATURE:

NAME/KEY: CDS

LOCATION: 149..3523

US-08-323-474-1

Alignment Scores:

Pred. No.: 1 Length: 4138
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.00% Indels: 0
DB: 1 Gaps: 0

US-09-397-967-16 (1-1099) x US-08-323-474-1 (1-4138)

OY 942 HlsarqaspleuAlaAlaargasnileuVal 952
Db 3032 CACAGGATCTGCTGCCAGAAACATTTAGTT 3064

RESULT 97

PCT-US93-06093-1

Sequence 1, Application PC/TUS9306093

GENERAL INFORMATION:

APPLICANT: Ziegler, Steven F.

TITLE OF INVENTION: NOVEL TYROSINE KINASE

NUMBER OF SEQUENCES: 3

CORRESPONDENCE ADDRESS:

ADDRESSEE: Immunex Corporation

STREET: 51 University Street

CITY: Seattle

STATE: Washington

COUNTRY: US

ZIP: 98101

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US93/06093

FILING DATE: 19930625

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/905,600

FILING DATE: 26-JUN-1992

ATTORNEY/AGENT INFORMATION:

NAME: Seese, Kathryn A.

REGISTRATION NUMBER: 32,172

REFERENCE/DOCKET NUMBER: 2609

TELEPHONE: (206) 587-0430

TELEFAX: (206) 233-0644

TELEX: 756822

INFORMATION FOR SEQ. ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 4138 base pairs

TYPE: NUCLEIC ACID

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA to mRNA

FEATURE:

NAME/KEY: CDS

LOCATION: 149..3523

PCT-US93-06093-1

Alignment Scores:
Pred. No.: 2.84 Length: 4138
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.00% Indels: 0
DB: 5 Gaps: 0

US-09-397-967-16 (1-1099) x PCT-US93-06093-1 (1-4138)

OY 942 HlsarqaspleuAlaAlaargasnileuVal 952
Db 3032 CACAGGATCTGCTGCCAGAAACATTTAGTT 3064

RESULT 98

US-08-442-248-1

Sequence 1, Application US/08442248

Patent No. 5759863

GENERAL INFORMATION:

APPLICANT: Caras, Ingrid W.

TITLE OF INVENTION: AL-1 Neurotrophic Factor

NUMBER OF SEQUENCES: 18

CORRESPONDENCE ADDRESS:

ADDRESSEE: Genentech, Inc.

STREET: 460 Point San Bruno Blvd

CITY: South San Francisco

STATE: California

COUNTRY: USA

ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patin (Genentech)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/442,248

FILING DATE: 15-MAY-1995

CLASSIFICATION: A35

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/330128

FILING DATE: 27-OCT-1994

ATTORNEY/AGENT INFORMATION:

NAME: Torchia, Timothy E.

REGISTRATION NUMBER: 36,700

REFERENCE/DOCKET NUMBER: 920C4

TELEPHONE: 415/225-8674

TELEFAX: 415/952-9881

TELEX: 910/371-7168

INFORMATION FOR SEQ. ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 4165 bases

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

US-08-442-248-1

Alignment Scores:
Pred. No.: 2.86 Length: 4165
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.00% Indels: 0
DB: 1 Gaps: 0

US-09-397-967-16 (1-1099) x US-08-442-248-1 (1-4165)

OY 941 ValhsarqaspleuAlaAlaargasnileu 951
Db 2602 GTGCACAGAGACCTGCTAGAACATCTTA 2634

RESULT 99

US-08-440-815-1

Sequence 1, Application US/08440815

Patent No. 5798448

GENERAL INFORMATION:

APPLICANT: Caras, Ingrid W.

TITLE OF INVENTION: AL-1 Neurotrophic Factor

NUMBER OF SEQUENCES: 18

CORRESPONDENCE ADDRESS:

ADDRESSEE: Genentech, Inc.

STREET: 460 Point San Bruno Blvd

CITY: South San Francisco

STATE: California

COUNTRY: USA

```

ZIP: 94080
COMPACT READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/440, 815
FILING DATE: 15-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/330128
FILING DATE: 27-OCT-1994
ATTORNEY/AGENT INFORMATION:
NAME: Torchia, Timothy E.
REGISTRATION NUMBER: 36,700
REFERENCE/DOCKET NUMBER: 920C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-8674
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4165 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

```

Alignment Scores:	
Pred. No.:	2.86
Score:	11.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	1.00%
DB:	1
Length:	4165
Matches:	11
Conservative:	0
Mismatches:	0
Indels:	0
Gaps:	0

US-09-397-967-16. (1-1099) x US-08-440-815-1 (1-4165)

QY	941	ValHisArgaspLeuAlaAlaArgsnlleLeu	951
Db	2602	GTCACAGAGACCTGTGCTAGAAACATCTTA	2634

RESULT 100

05 00 400 472 1
; Sequence 1, Application US/08486449

; Patent No. 6280732

GENERAL INFORMATION:

APPLICANT: Wincslaw, John W.

TITLE OF INVENTION: AL-1 Neurotrophic Factor

NUMBER OF SEQUENCES: 18

CORRESPONDENCE ADDRESS:

ADDRESSEE: Genentech, Inc.
460 Point San Bruno Blvd
Sunnyvale, CA 94086

STREET: 480 Point San Diego
CITY: South San Francisco

STATE: California

COUNTRY: USA

ZIP: 94080

```

; COMPUTER READABLE FORM:
;
; 3 5 inch 1 44 MB floppy disk
;

```

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: patin (Genentech)

CURRENT APPLICATION DATA: TC 103/485 440

APPLICATION NUMBER: US/0
05-TIN-1995

FILED DATE: 06-JUN-2006
CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/330128

FILING DATE: 27-OCT-1994

ATTORNEY/AGENT INFORMATION:

```

1      REGISTRATION NUMBER: 36,700
2      REFERENCE/DOCKET NUMBER: P0920P1
3      TELECOMMUNICATION INFORMATION:
4      TELEPHONE: 415/225-8674
5      TELEFAX: 415/952-9881
6      TELEX: 910/371-7168
7      INFORMATION FOR SEQ ID NO: 1:
8      SEQUENCE CHARACTERISTICS:
9      LENGTH: 4165 bases
10     TYPE: nucleic acid
11     STRANDEDNESS: single
12     TOPOLOGY: linear
13
14     US-08-486-449-1
15
16     Alignment Scores:
17     Pred. No.:
18     Score: 2.86
19     Percent Similarity: 11.00%
20     Best Local Similarity: 100.00%
21     Query Match: 1.00%
22
23     DB: 4
24
25     US-09-397-967-16 (1-1099) x US-08-486-449-1 (1-4165)
26
27     QY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeu 951
28     |||||||
29     Db 2602 GTGCACAGACACCTTGCTGCTAAGAACATCTTA 2634
30
31     Search completed: April 28, 2003, 21:44:09
32     Job time : 357 secs

```

```
US-09-397-967-16 (1-1099) x US-08-466-449-1 (1-4165)  
QY      941 ValHisArgSpLenAlaIaaIArgAsnLeu 951  
          |||||  
Db       2602 GTGACACAGAGACTTGGTGCTAGAACAATCTTA 2634
```

Search completed: April 28, 2003, 21:44:09
Job time : 357 secs